Available online at www.sciencedirect.com
ScienceDirect
Journal of Differential Equations 301 (2021) 202–235
Journal of Differential Equations
www.elsevier.com/locate/jde

Averaging principle for fast-slow system driven by mixed fractional Brownian rough path

Bin Pei [a,b], Yuzuru Inahama [b], Yong Xu [a,*]

[a] *School of Mathematics and Statistics, Northwestern Polytechnical University, 127 West Youyi Rd., Beilin District, Xi'an, Shaanxi, 710072, China*

[b] *Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan*

Received 6 December 2020; revised 5 May 2021; accepted 3 August 2021
Available online 26 August 2021

Abstract

This paper is devoted to studying the averaging principle for a fast-slow system of rough differential equations driven by mixed fractional Brownian rough path. The fast component is driven by Brownian motion, while the slow component is driven by fractional Brownian motion with Hurst index H $(1/3 < H \leq 1/2)$. Combining the fractional calculus approach to rough path theory and Khasminskii's classical time discretization method, we prove that the slow component strongly converges to the solution of the corresponding averaged equation in the L^1-sense. The averaging principle for a fast-slow system in the framework of rough path theory seems new.

MSC: 60G22; 60H10; 34C29

Keywords: Averaging principle; Fast-slow system; Mixed fractional Brownian rough path; Fractional calculus approach

* Corresponding author.
E-mail addresses: binpei@nwpu.edu.cn (B. Pei), inahama@math.kyushu-u.ac.jp (Y. Inahama), hsux3@nwpu.edu.cn (Y. Xu).

https://doi.org/10.1016/j.jde.2021.08.006

1. Introduction

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a complete probability space, $W = (W_t)_{t\geq 0}$ be a standard d'-dimensional Brownian motion (Bm), $B = (B_t)_{t\geq 0}$ be a d-dimensional fractional Brownian motion (fBm) with Hurst index $H \in (\frac{1}{3}, \frac{1}{2}]$, that is a collection of centered, independent Gaussian processes, independent of W as well, with covariance function

$$R_H(t,s) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right) I_d, \quad s,t \geq 0,$$

where I_d is the identity matrix of size d. The Kolmogorov theorem entails that fBm has a modification with β-Hölder sample paths for any $\beta < H$. For each $t \geq 0$, we denote by $\mathcal{F}_t$ the σ-field generated by the random variables $\{B_s, W_s, s \in [0,t]\}$ and all P-null sets. The expectation with respect to P is denoted by $\mathbb{E}$. In addition to the natural filtration $\{\mathcal{F}_t, t \geq 0\}$, we will consider a larger filtration $\{\mathcal{G}_t, t \geq 0\}$ such that $\{\mathcal{G}_t\}$ is right-continuous and $\{\mathcal{G}_0\}$ contains the P-null sets, so that B are $\mathcal{G}_0$-measurable, and W is a $\{\mathcal{G}_t\}$-Brownian motion.

In what follows, we will denote by $C_b^k(\mathbb{R}^m; \mathbb{R}^n)$ the set of functions $f : \mathbb{R}^m \to \mathbb{R}^n$ which are bounded, k-times continuously differentiable with bounded derivatives of order up to k (in symbols $\nabla f, \nabla^2 f, \ldots, \nabla^k f$). We denote by $C_b^{k,\lambda}(\mathbb{R}^m; \mathbb{R}^n)$ the set of $f \in C_b^k(\mathbb{R}^m; \mathbb{R}^n)$ whose kth derivative is uniformly Hölder continuous of order $\lambda \in (0,1]$. The time interval will be $[0,T]$ for arbitrary $T > 0$. The transpose of a vector A will be denoted by A^{T}.

We firstly consider the differential equation driven by both Bm and fBm of the type

$$u_t = u_0 + \int_0^t a(u_s)\mathrm{d}s + \int_0^t b(u_s)\mathrm{d}B_s + \int_0^t c(u_s)\mathrm{d}W_s, \tag{1.1}$$

with $u_0 \in \mathbb{R}^e$ which is arbitrary and non-random but fixed. Since the fBm is neither Markov nor semimartingale if $H \neq \frac{1}{2}$, we cannot use the classical Itô theory to solve (1.1) unless $b \equiv 0$. Lyons [21] made a breakthrough by inventing rough path theory, which enabled us to do pathwise study of stochastic differential equations (SDEs) as above. We will show that (1.1) can be understood as a rough differential equation (RDE) and possesses a unique global solution if the coefficients $a : \mathbb{R}^e \to \mathbb{R}^e$, $b : \mathbb{R}^e \to \mathbb{R}^e \otimes \mathbb{R}^d$ and $c : \mathbb{R}^e \to \mathbb{R}^e \otimes \mathbb{R}^{d'}$ satisfy suitable regularity assumptions.

The driving rough path of RDE (1.1) is a natural rough path lift of $(B_t, W_t)_{0\leq t\leq T}$, which is formally given by $\mathbf{Z} = (Z, Z^2)$, where

$$Z_t = (B_t, W_t)^{\mathrm{T}} \quad \text{and} \quad Z_{st}^2 = \begin{pmatrix} B_{st}^2 & \int_s^t (B_u - B_s)\otimes \mathrm{d}W_u \\ \int_s^t (W_u - W_s)\otimes \mathrm{d}B_u & W_{st}^2 \end{pmatrix}. \tag{1.2}$$

For every $\beta \in (\frac{1}{3}, H)$, this lift $\mathbf{Z}$ exists and is a β-Hölder (weakly) geometric rough path almost surely (see Section 3 for details). We call it mixed fractional Brownian rough path. Since the W-component of $\mathbf{Z}$ is Stratonovich-type Brownian rough path, the last term in (1.1) is something like a Stratonovich integral.

In the case $H \in (\frac{1}{2}, 1)$, it is well-known that Young integral, which is essentially a generalized Riemann-Stieltjes integral, could be a good choice to give meaning to the integral with respect to fBm [26,37]. When $c \equiv 0$ in (1.1), the theory of SDEs driven only by fBm with $H \in (\frac{1}{2}, 1)$ was initiated by Lyons [20] and has been well developed by many authors especially on the existence and uniqueness of pathwise solutions. For example, using the fractional calculus introduced by Zähle [37], Nualart and Răşcanu [26] derive very weak conditions for the general case where a, b are functions of (t, u_t), in particular b need to be only C^1 with bounded and Hölder continuous first derivative, to ensure the existence and uniqueness of the solution in the space of Hölder continuous functions. When c does not vanish identically, there are only a few results devoted to such mixed equations. The main difficulty when considering (1.1) lies in the fact that both stochastic integrals are dealt in different ways. The integral with respect to the Bm is understood as an Itô integral, while the integral with respect to the fBm has to be understood in the pathwise Young sense. Kubilius [18] studies SDEs driven by both fBm and Bm, in the one-dimensional case, with no drift term. Guerra and Nualart [13] combine the pathwise approach (generalized Riemann-Stieltjes integral) with the Itô stochastic calculus to prove an existence and uniqueness theorem for multidimensional, time-dependent SDEs driven simultaneously by a multidimensional fBm with Hurst parameter $H > 1/2$ and a multidimensional Bm.

Let us get back to the case $H \in (\frac{1}{3}, \frac{1}{2}]$. When $c \equiv 0$ in (1.1), the existence and uniqueness result was generalized by Lyons' seminal paper [21]. In this paper he established rough path theory to define the integral with respect to β-Hölder rough path ($\frac{1}{3} < \beta < \frac{1}{2}$). This theory is basically deterministic (see several monographs [8,9,22,23]). Coutin and Qian [4] proved that fBm admits a natural rough path lift in $1/\beta$-variation rough path topology, which was later improved to β-Hölder rough path topology ($\frac{1}{3} < \beta < H$). These results enabled us to study (1.1) via rough paths when $c \equiv 0$.

There are other formulations of rough path theory. Gubinelli [12] established an alternative theory of controlled rough paths to generalize the concept of integration and differential equations with Hölder exponent greater than $\frac{1}{3}$. Following Zähle [37], Hu and Nualart [15] developed another approach to rough path theory by using fractional calculus.

Like the case of usual SDEs, the condition on the drift a can be weaker than one on the diffusion coefficient b. A well-known result states that a unique global solution exists when $a \in C_b^1$ and $b \in C_b^3$ (see [10] for instance). Riedel and Scheutzow [31] solved RDEs with unbounded drift term. In this work, a is allowed to grow at most linearly, while $b \in C_b^4$ (see Proposition 1.1 below).

When c does not vanish identically, however, much less is known about (1.1). The most relevant result in our case is Diehl, Oberhauser and Riedel [5], in which the authors gave a meaning to differential equations driven by a deterministic rough path and Brownian rough path. In contrast to the RDE in [5], the trajectories of B and W in (1.1) are both stochastic.

Now, we summarize the existence and uniqueness result for (1.1). Needless to say, the unique solution does not depend on the choice of $\beta \in (\frac{1}{3}, H)$.

Proposition 1.1. *Let $\frac{1}{3} < \beta < H \leq \frac{1}{2}$ and write $b = (b_1, \ldots, b_d)$ and $c = (c_1, \ldots, c_{d'})$. Assume either one of the following two conditions on the coefficients of RDE (1.1).*

1. *a is a locally Lipschitz continuous vector field with at most linear growth on $\mathbb{R}^e$ and $b_i, c_j \in C_b^4(\mathbb{R}^e, \mathbb{R}^e)$ $(1 \leq i \leq d, 1 \leq j \leq d')$.*
2. *$a \in C_b^1(\mathbb{R}^e, \mathbb{R}^e)$ and $b_i, c_j \in C_b^3(\mathbb{R}^e, \mathbb{R}^e)$ $(1 \leq i \leq d, 1 \leq j \leq d')$.*

Then, RDE (1.1) possesses a unique global solution in the framework of β-Hölder rough path theory.

Next, we will deal with a fast-slow system of RDEs driven by mixed fractional Brownian rough path with Hurst index $H \in (\frac{1}{3}, \frac{1}{2}]$ of the type

$$\begin{cases} X_t^\varepsilon = X_0^\varepsilon + \int_0^t f\left(X_s^\varepsilon, Y_s^\varepsilon\right) \mathrm{d}s + \int_0^t \sigma\left(X_s^\varepsilon\right) \mathrm{d}B_s, \\ Y_t^\varepsilon = Y_0^\varepsilon + \frac{1}{\varepsilon}\int_0^t g\left(X_s^\varepsilon, Y_s^\varepsilon\right) \mathrm{d}s + \frac{1}{\sqrt{\varepsilon}}\int_0^t h\left(X_s^\varepsilon, Y_s^\varepsilon\right) \mathrm{d}W_s, \end{cases} \tag{1.3}$$

where $X_0^\varepsilon = X_0 \in \mathbb{R}^m$, $Y_0^\varepsilon = Y_0 \in \mathbb{R}^n$ are arbitrary and non-random but fixed, while ε is a small positive parameter. The coefficients $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^m \to \mathbb{R}^m \otimes \mathbb{R}^d$, $h : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n \otimes \mathbb{R}^{d'}$ satisfy the following regularity assumptions:

(H1) f is a Lipschitz continuous vector field with at most linear growth, $\sigma \in C_b^4(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^d)$.
(H2) g is a Lipschitz continuous vector field with at most linear growth, $h \in C_b^4(\mathbb{R}^m \times \mathbb{R}^n; \mathbb{R}^n \otimes \mathbb{R}^{d'})$.

The system of RDEs (1.3) is a special case of RDE (1.1) and therefore possesses a unique global solution under (H1) and (H2) due to Proposition 1.1.

Let us summarize some basic results on the stochastic averaging principle for this kind of slow-fast systems, which can be traced back to the work of Khasminskii [17], see e.g. [6,7,19,29,30,33–36] and the references therein. In order to obtain the strong convergence, it is known that the diffusion coefficient σ in (1.3) should not depend on the fast variable Y^ε (see e.g. [11]). The corresponding literature in the case of perturbation by multiplicative fractional Brownian noise is quite sparse. It is worth mentioning that Hairer and Li [14] considered a slow-fast system where the slow component is driven by fBm and proved the convergence to the averaged solution takes place in probability. The most relevant result in our case is the recent work [27,32] which answered affirmatively that an averaging principle still holds for fast-slow mixed SDEs if disturbances involve both Bm and fBm with $H \in (\frac{1}{2}, 1)$ in the mean square sense. Using the generalized Riemann-Stieltjes integral, an averaging principle in the mean square sense for stochastic partial differential equations driven by fBm subject to an additional fast-varying diffusion process was established in [28]. We point out that all the aforementioned papers concerning fBm assume $H > \frac{1}{2}$.

Therefore, it is quite natural to extend the averaging result to the case $H \in (\frac{1}{3}, \frac{1}{2}]$. Compared to the known results, the main difficulty here is how to deal with the mixed fractional Brownian rough path with Hurst index $H \in (\frac{1}{3}, \frac{1}{2}]$. We will mainly use Hu and Nualart's fractional calculus approach to prove the averaging principle (see Theorem 1.2 below).

Now, following the averaging theory inspired by Khasminskii in [17], we define the averaged RDE as follows:

$$\bar{X}_t = \bar{X}_0 + \int_0^t \bar{f}(\bar{X}_s)\mathrm{d}s + \int_0^t \sigma(\bar{X}_s)\mathrm{d}B_s, \tag{1.4}$$

with $\bar{X}_0 = X_0$, where we set

$$\bar{f}(\xi) = \int_{\mathbb{R}^n} f(\xi, \phi)\mu^{\xi}(\mathrm{d}\phi), \quad \xi \in \mathbb{R}^m, \tag{1.5}$$

for a unique invariant probability measure μ^{ξ} with respect to the following frozen SDE under condition (H4) below:

$$\mathrm{d}Y_t^{\xi,\phi} = \tilde{g}(\xi, Y_t^{\xi,\phi})\mathrm{d}t + h(\xi, Y_t^{\xi,\phi})\mathrm{d}^{\mathrm{I}} W_t, \quad Y_0^{\xi,\phi} = \phi \in \mathbb{R}^n. \tag{1.6}$$

Here, $\int \cdots \mathrm{d}^{\mathrm{I}} W$ stands for the usual Itô integral and

$$\tilde{g}(\xi, \phi) = g(\xi, \phi) + \frac{1}{2}\sum_{\bar{l}=1}^{n}\sum_{\bar{j}=1}^{d'} \mathcal{D}_h^{(\bar{j})} h_{\bar{l},\bar{j}}(\xi, \phi), \quad \mathcal{D}_h^{(\bar{j})} = \sum_{\bar{l}=1}^{n} h_{\bar{l},\bar{j}}(\cdot, \cdot)\partial_{\phi_{\bar{l}}}.$$

To establish the averaging principle for (1.3), we set the following hypotheses:

(H3) $f \in C_b^1(\mathbb{R}^m \times \mathbb{R}^n; \mathbb{R}^m)$.
(H4) There exist $L > 0$, $\beta_i > 0$, $i = 1, 2$, such that

$$2\langle \phi - \tilde{\phi}, \tilde{g}(\xi, \phi) - \tilde{g}(\xi, \tilde{\phi})\rangle + |h(\xi, \phi) - h(\xi, \tilde{\phi})|^2 \le -\beta_1|\phi - \tilde{\phi}|^2,$$
$$2\langle \phi, \tilde{g}(\xi, \phi)\rangle + |h(\xi, \phi)|^2 \le -\beta_2|\phi|^2 + L|\xi|^2 + L$$

for any $\xi \in \mathbb{R}^m$ and $\phi, \tilde{\phi} \in \mathbb{R}^n$.

Now, we present our main result of averaging principle in the L^1-sense. To our knowledge, this is the first result that proves the averaging principle for a fast-slow system in the framework of rough path theory.

Theorem 1.2. *Let $\frac{1}{3} < H \le \frac{1}{2}$ and $\mathbf{Z}$ be the natural rough path lift of $(B_t, W_t)_{t\in[0,T]}$ as in (1.2). Assume that f, σ, g, h satisfy* (H1)-(H4). *Then, we have*

$$\lim_{\varepsilon \to 0} \mathbb{E}[\|X^{\varepsilon} - \bar{X}\|_{\infty}] = 0.$$

Here, $\|\cdot\|_{\infty}$ denotes the supremum norm over $[0, T]$ and X^{ε} and $\bar{X}$ denote the first level paths of the slow component of (1.3) and (1.4), respectively.

Remark 1.3. A simple example that satisfies the (H4) is $g(\xi, \phi) = \xi - 8\phi$ and $h(\xi, \phi) = \sin\xi + \sin\phi$ when $d = d' = m = n = 1$. Another example is as follows. Let $g(\xi, \phi) = -A(\xi)\phi$, where A is a bounded, positive, C_b^1-function in ξ, which is also bounded away from zero. If $\|h\|_{\infty} + \|\nabla_{\phi} h\|_{\infty} + \|\nabla_{\phi}^2 h\|_{\infty}$ is sufficiently small, then these g and h satisfy (H4). Here, ∇_{ϕ} stands for the (partial) gradient with respect to ϕ and $\|\cdot\|_{\infty}$ denotes the supremum norm.

The rest of the paper is organized as follows. Section 2 presents some notations and the pathwise approach based on the techniques of the fractional calculus and rough path theory. The existence and uniqueness theorems to (1.1) and (1.3) are proved in Section 3. Section 4 is devoted to proving Theorem 1.2, that is, the averaging principle for the fast-slow system driven by mixed fractional Brownian rough path with Hurst index $H \in (\frac{1}{3}, \frac{1}{2}]$.

Throughout this paper, K and K_* denote certain positive constants that may vary from line to line. K_* is used to emphasize that the constant depends on the corresponding parameter $*$, which is one or more than one parameter.

2. Preliminaries

Let $|\cdot|$ stand for an absolute value of a real number, the Euclidean norm of a finite dimensional vector or of a matrix, $\langle \cdot, \cdot \rangle$ be the Euclidean inner product. Set $\Delta_{a,b} = \{(s,t) \mid a \le s \le t \le b\}$. Moreover, for a function $f : [a,b] \to \mathbb{R}^m$ we define the following seminorms:

$$\|f\|_{\infty,[a,b]} = \sup_{t\in[a,b]} |f(t)|, \quad \|f\|_{\gamma,[a,b]} = \sup_{(s,t)\in\Delta_{a,b}} \frac{|f(t)-f(s)|}{|t-s|^{\gamma}},$$

where $\gamma \in (0,1]$ and also use the convention $0/0 \triangleq 0$. We will study continuous $\mathbb{R}^m$-valued paths on some interval $[a,b]$, and we denote the space of such functions by $C([a,b];\mathbb{R}^m)$. We denote by $C^{\gamma}([a,b];\mathbb{R}^m)$, the space of γ-Hölder continuous functions on some interval $[a,b]$ with values in $\mathbb{R}^m$ and set $\|f\|_{\gamma} := \|f\|_{\gamma,[0,T]}$, $\|f\|_{\infty} := \|f\|_{\infty,[0,T]}$ and $\Delta := \Delta_{0,T}$ for shortness. Next, for a function $v : \Delta_{a,b} \to \mathbb{R}^m$, that vanishes on the diagonal, that is, $v(t,t) = 0$ for $t \in [a,b]$, we set

$$\|v\|_{\gamma,\Delta_{a,b}} = \sup_{(s,t)\in\Delta_{a,b}} \frac{|v(s,t)|}{|t-s|^{\gamma}}.$$

The set of such functions with a finite norm $\|v\|_{\gamma,\Delta_{a,b}}$ is denoted by $C_2^{\gamma}(\Delta_{a,b};\mathbb{R}^m)$,

$$C_2^{\gamma}(\Delta_{a,b};\mathbb{R}^m) = \{v : \Delta_{a,b} \to \mathbb{R}^m \mid v(t,t) = 0, t \in [a,b], \|v\|_{\gamma,\Delta_{a,b}} < \infty\}.$$

The purpose of this paper is to use the pathwise approach including rough path analysis and an approach via fractional calculus to the stochastic calculus with respect to Bm and fBm. To prove Proposition 1.1, we mainly use the rough path analysis developed in Lyons [20]. To establish the averaging principle (see Theorem 1.2), we combine the rough path analysis and the pathwise approach via fractional calculus inspired by the work of Hu and Nualart [15].

We firstly recall the pathwise approach based on the techniques of the fractional calculus in forthcoming Section 2.1 (see Hu and Nualart [15] and Definition 2.2 below). Then, following the ideas in [12] and [20], we provide an explicit formula for integrals in the rough path sense (see Gubinelli in [12] and (2.11) in Section 2.2).

2.1. Integrals along rough paths via fractional calculus

Without approximation by Riemann sums, a pathwise approach to study integration of vector-valued Hölder continuous rough functions using fractional calculus was firstly developed by Hu and Nualart [15]. Later, Ito [16] showed that the integral defined in Definition 2.2 below (see also Definition 3.2 in [15]) using fractional calculus should be consistent with that obtained by the usual integration in rough path analysis, that will be given by the limit of the compensated Riemann-Stieltjes sums in forthcoming (2.11) of next subsection. In this subsection, we follow the notations proposed by Hu and Nualart [15].

This subsection aims to recall the integrals along rough paths via fractional calculus, that is giving meaning to the following integral

$$\int_0^T \sigma(x_r)\mathrm{d}\omega_r = \sum_{j=1}^{d} \int_0^T \sigma_j(x_r)\mathrm{d}\omega_r^j, \tag{2.1}$$

using fractional calculus, where $\omega \in C^\beta([0,T];\mathbb{R}^d)$ with $\beta \in (\frac{1}{3}, \frac{1}{2})$. Moreover, for a given $\omega \in C^\beta([0,T];\mathbb{R}^d)$, consider v to be an element of $C_2^{2\beta}(\Delta;\mathbb{R}^m \otimes \mathbb{R}^d)$ and assume that the triplets (x,ω,v) satisfy Chen's relation: for all $0 \le s \le \tau \le t \le T$ it holds that

$$v_{s\tau} + v_{\tau t} + (x_\tau - x_s) \otimes (\omega_t - \omega_\tau) = v_{st}, \tag{2.2}$$

where $\otimes$ denotes tensor.

Definition 2.1.

(1) For a given $\omega \in C^\beta([0,T];\mathbb{R}^d)$, we denote by $M_{m,d}^\beta$ the space consisting of triplets $(x,\omega,v) \in C^\beta([0,T];\mathbb{R}^m) \times C^\beta([0,T];\mathbb{R}^d) \times C_2^{2\beta}(\Delta;\mathbb{R}^m \otimes \mathbb{R}^d)$ such that (2.2) holds.

(2) (ω,ω^2) is called a β-Hölder rough path (over $\mathbb{R}^d$) if $(\omega,\omega,\omega^2) \in M_{d,d}^\beta$. Moreover, define the space of β-Hölder rough paths (over $\mathbb{R}^d$) in symbols $\mathscr{C}^\beta([0,T];\mathbb{R}^d)$ and the space of weakly geometric β-Hölder rough paths in symbols $\mathscr{C}_g^\beta([0,T];\mathbb{R}^d)$ by stipulating that $(\omega,\omega^2) \in \mathscr{C}_g^\beta([0,T];\mathbb{R}^d)$ if and only if $(\omega,\omega^2) \in \mathscr{C}^\beta([0,T];\mathbb{R}^d)$ and

$$\mathrm{Sym}(\omega_{st}^2) = \frac{1}{2}(\omega_t - \omega_s) \otimes (\omega_t - \omega_s),$$

for every s,t.

To proceed, for $\alpha \in (0,1)$, define the following fractional derivatives

$$D_{s+}^\alpha f[r] = \frac{1}{\Gamma(1-\alpha)}\Big(\frac{f(r)}{(r-s)^\alpha} + \alpha \int_s^r \frac{f(r)-f(\theta)}{(r-\theta)^{\alpha+1}}\mathrm{d}\theta\Big),$$

$$D_{t-}^{1-\alpha} g_{t-}[r] = \frac{(-1)^{1-\alpha}}{\Gamma(\alpha)}\Big(\frac{g(r)-g(t)}{(t-r)^{1-\alpha}} + (1-\alpha)\int_r^t \frac{g(r)-g(\theta)}{(\theta-r)^{2-\alpha}}\mathrm{d}\theta\Big),$$

where $g_{t-} = g(\cdot) - g(t)$.

Then, we are now ready to define the integral $\int_a^b \sigma(x_r)\mathrm{d}\omega_r$.

Definition 2.2 *(cf. [15, Definition 3.2]).* Let $\beta \in (\frac{1}{3}, \frac{1}{2})$, $(x, \omega, v) \in M^{\beta}_{m,d}$ and $\sigma \in C^{1,\lambda}(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^d)$ with $(2+\lambda)\beta > 1$. Fix $\alpha \in (0, 1)$ such that $1 - \beta < \alpha < 2\beta$, and $\alpha < \frac{\lambda\beta+1}{2}$. Then, for any $(a, b) \in \Delta$ we define

$$\int_a^b \sigma(x_r)\mathrm{d}\omega_r = (-1)^{\alpha} \int_a^b \hat{D}^{\alpha}_{a+}\sigma(x)[r] D^{1-\alpha}_{b-}\omega_{b-}[r]\mathrm{d}r$$
$$- (-1)^{2\alpha-1} \int_a^b D^{2\alpha-1}_{a+}\nabla\sigma(x)[r] D^{1-\alpha}_{b-}\mathcal{D}^{1-\alpha}_{b-}v[r]\mathrm{d}r.$$

Here, for $r \in (a, b)$,

$$\hat{D}^{\alpha}_{a+}\sigma(x)[r] = \frac{1}{\Gamma(1-\alpha)}\left(\frac{\sigma(x_r)}{(r-a)^{\alpha}} + \alpha \int_a^r \frac{\sigma(x_r) - \sigma(x_\theta) - \nabla\sigma(x_\theta)(x_r - x_\theta)}{(r-\theta)^{\alpha+1}}\mathrm{d}\theta\right)$$

is the compensated fractional derivative and

$$\mathcal{D}^{1-\alpha}_{b-}v[r] = \frac{(-1)^{1-\alpha}}{\Gamma(\alpha)}\left(\frac{v_{rb}}{(b-r)^{1-\alpha}} + (1-\alpha)\int_r^b \frac{v_{rs}}{(s-r)^{2-\alpha}}\mathrm{d}s\right)$$

is the extension of the fractional derivative of v.

Notice that under the constraints $1 - \beta < \alpha < 2\beta$ and $\alpha < \frac{\lambda\beta+1}{2}$, it is easy to prove that the fractional derivatives $D^{1-\alpha}_{b-}\omega_{b-}[r]$ and $D^{1-\alpha}_{b-}\mathcal{D}^{1-\alpha}_{b-}v[r]$ are well defined because the functions ω and $\mathcal{D}^{1-\alpha}_{b-}v$ are β-Hölder continuous. Because there exists a constant $K > 0$ such that for all $r, \theta \in [a, b], \theta < r$, we have

$$|\sigma(x_r) - \sigma(x_\theta) - \nabla\sigma(x_\theta)(x_r - x_\theta)| \leq K|r-\theta|^{(1+\lambda)\beta},$$
$$|\nabla\sigma(x_r) - \nabla\sigma(x_\theta)| \leq K|r-\theta|^{\beta\lambda},$$

then the derivatives $\hat{D}^{\alpha}_{a+}\sigma(x)[r]$ and $D^{2\alpha-1}_{a+}\nabla\sigma(x)[r]$ are also well-defined. More details can be found e.g. in [15, p. 2694].

Now, given a continuous path $y \in C([0, T]; \mathbb{R}^n)$, we aim to solve the RDE with drift term driven by a β-Hölder rough path (ω, ω^2)

$$x_t = x_0 + \int_0^t f(x_s, y_s)\mathrm{d}s + \int_0^t \sigma(x_s)\mathrm{d}\omega_s, \tag{2.3}$$

where $f \in C^1_b(\mathbb{R}^m \times \mathbb{R}^n; \mathbb{R}^m)$, $\sigma \in C^3_b(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^d)$.

The main idea to solve (2.3) inspired by Hu and Nualart in [15, Section 4] and Garrido-Atienza and Schmalfuss [10, Theorem 4] is to write a system of three components for the enlarged unknown (x, ω, v). According to Definition 2.2, the first component is just (2.3) itself, where the right-hand side is a function of (x, ω, v), i.e.

$$x_t = x_0 + \int_0^t f(x_s, y_s)\mathrm{d}s + (-1)^\alpha \int_0^t \hat{D}_{0+}^\alpha \sigma(x)[r] D_{t-}^{1-\alpha} \omega_{t-}[r]\mathrm{d}r$$
$$-(-1)^{2\alpha-1} \int_0^t D_{0+}^{2\alpha-1} \nabla\sigma(x)[r] D_{t-}^{1-\alpha} \mathcal{D}_{t-}^{1-\alpha} v[r]\mathrm{d}r.$$

The second component is

$$v_{st} = \int_s^t \int_s^r f(x_q, y_q)\mathrm{d}q \otimes \mathrm{d}\omega_r - \int_s^t \sigma(x_r)\mathrm{d}\omega^2_{\cdot,t}(r). \tag{2.4}$$

Note that the second term on the right-hand side of (2.4) is a functional of (x, ω^2, w) again by Definition 2.2 (w will be given later), i.e.

$$\int_s^t \sigma(x_r)\mathrm{d}\omega^2_{\cdot,t}(r) = (-1)^\alpha \int_s^t \hat{D}_{s+}^\alpha \sigma(x)[r] D_{t-}^{1-\alpha} \omega^2_{\cdot,t-}[r]\mathrm{d}r$$
$$-(-1)^{2\alpha-1} \int_s^t D_{s+}^{2\alpha-1} \nabla\sigma(x)[r] D_{t-}^{1-\alpha} \mathcal{D}_{t-}^{1-\alpha} w_{t,\cdot,t}[r]\mathrm{d}r. \tag{2.5}$$

The third component is defined by writing w as a functional of (x, ω, v, ω^2) (see [15, (3.26)]) as follows, for $s \le q \le t$

$$\begin{aligned} w_{t,s,q} = &-\frac{(-1)^{2\alpha-1}}{\Gamma(2-2\alpha)} \int_s^t \left[\frac{x_r - x_s}{(r-s)^{2\alpha-1}} + (2\alpha-1)\int_s^r \frac{x_r - x_\theta}{(r-\theta)^{2\alpha}}\mathrm{d}\theta \right] \otimes D_{q-}^{1-\alpha}\mathcal{D}_{q-}^{1-\alpha}\omega^2[r]\mathrm{d}r \\ &+\frac{(-1)^{2\alpha-1}}{\Gamma(2-2\alpha)} \int_s^q \left[\frac{\omega_t - \omega_r}{(r-s)^{2\alpha-1}} + (2\alpha-1)\int_s^r \frac{\omega_\theta - \omega_r}{(r-\theta)^{2\alpha}}\mathrm{d}\theta \right] \otimes D_{q-}^{1-\alpha}\mathcal{D}_{q-}^{1-\alpha} v[r]dr \\ &-\frac{(-1)^\alpha}{\Gamma(1-\alpha)} \int_s^q \left(\frac{(x_r - x_s)\otimes(\omega_t - \omega_r)}{(r-s)^\alpha} + \alpha \int_s^r \frac{(x_\theta - x_r)\otimes(\omega_r - \omega_\theta)}{(r-\theta)^{\alpha+1}}\mathrm{d}\theta \right) \\ &\otimes D_{q-}^{1-\alpha}\omega_{q-}[r]\mathrm{d}r. \end{aligned} \tag{2.6}$$

Remark 2.3. It should be noted that the sign in front of the second term on the right-hand side of (2.4) is negative. This definition of RDEs was first given in [15, p. 2701, Eq. (4.2)] when $f \equiv 0$.

However, a negative sign is missing there (and in many other subsequent works). Concerning this, the right side of (2.6) and that of [15, p. 2701, Eq. (4.5)] have the opposite signs. Fortunately, since what is actually computed is the norm of these terms, all the results in [15] remain valid.

By a slight generalization of Theorem 4 in [10], a solution of (2.3) is defined to be an element of $M_{m,d}^{\beta}$ when (2.3)-(2.6) have a solution and given a continuous path $y \in C([0,T];\mathbb{R}^n)$ and a rough path $(\omega,\omega^2) \in \mathscr{C}^{\beta}([0,T];\mathbb{R}^d)$ with $\beta \in (\frac{1}{3},\frac{1}{2})$, (2.3) has a unique global solution under the conditions $f \in C_b^1(\mathbb{R}^m \times \mathbb{R}^n;\mathbb{R}^m)$, $\sigma \in C_b^3(\mathbb{R}^m;\mathbb{R}^m \otimes \mathbb{R}^d)$, see Garrido-Atienza and Schmalfuss [10, Theorem 4] for example.

Based on a slight generalization of these results in Hu and Nualart [15] and Besalú, Binotto and Rovira [1], it is not difficult to provide quantitative estimates for the integration operator in (2.3) and (2.4). Let us recall two propositions from [1,2,15]. On $M_{m,d}^{\beta}$, we introduce the following functionals for any $(s,t) \in \Delta$:

$$\Phi_{\beta,[s,t]}(x,\omega,v) = \|v\|_{2\beta,\Delta_{s,t}} + \|x\|_{\beta,[s,t]}\|\omega\|_{\beta,[s,t]}, \tag{2.7}$$

and

$$\Phi_{\beta,[s,t]}(x,\omega,v,\omega^2) = \|\omega\|_{\beta,[s,t]}^2\|x\|_{\beta,[s,t]} + \|\omega\|_{\beta,[s,t]}\|v\|_{2\beta,\Delta_{s,t}} + \|x\|_{\beta,[s,t]}\|\omega^2\|_{2\beta,\Delta_{s,t}}. \tag{2.8}$$

We have the following estimates:

Proposition 2.4 *(cf. [15, Proposition 3.4]). Let $(x,\omega,v) \in M_{m,d}^{\beta}$ and $\sigma \in C_b^{1,\lambda}(\mathbb{R}^m;\mathbb{R}^m \otimes \mathbb{R}^d)$ with $(2+\lambda)\beta > 1$. Then, for any $(a,b) \in \Delta$, we have*

$$\left|\int_a^b \sigma(x_r)\mathrm{d}\omega_r\right| \le K|\sigma(x_a)|\|\omega\|_{\beta,[a,b]}(b-a)^{\beta}$$
$$+K\Phi_{\beta,[a,b]}(x,\omega,v)(\|\nabla\sigma\|_\infty + \|\nabla\sigma\|_\lambda\|x\|_{\beta,[a,b]}^{\lambda}(b-a)^{\lambda\beta})(b-a)^{2\beta},$$

where $\Phi_{\beta,[a,b]}(x,\omega,v)$ is defined in (2.7).

Proposition 2.5 *(cf. [15, Proposition 3.9]). Let $(x,\omega,v) \in M_{m,d}^{\beta}$, $(\omega,\omega^2) \in \mathscr{C}^{\beta}([0,T];\mathbb{R}^d)$ and $\sigma \in C_b^{1,\lambda}(\mathbb{R}^m;\mathbb{R}^m \otimes \mathbb{R}^d)$ with $(2+\lambda)\beta > 1$. Then, for any $(a,b) \in \Delta$, we have*

$$\left|\int_a^b \sigma(x_r)\mathrm{d}\omega_{\cdot,b}^2(r)\right| \le K|\sigma(x_a)|\Phi_{\beta,[a,b]}(\omega,\omega,\omega^2)(b-a)^{2\beta}$$
$$+K(\|\nabla\sigma\|_\infty + \|\nabla\sigma\|_\lambda\|x\|_{\beta,[a,b]}^{\lambda}(b-a)^{\lambda\beta})\Phi_{\beta,[a,b]}(x,\omega,v,\omega^2)(b-a)^{3\beta},$$

where $\Phi_{\beta,[a,b]}(x,\omega,v,\omega^2)$ is defined in (2.8).

Given $(\omega, \omega^2) \in \mathscr{C}^\beta([0,T];\mathbb{R}^d)$, we write

$$\Lambda_\omega = 1 + \|\omega\|_\beta^2 + \|\omega^2\|_{2\beta}.$$

In [1,3], Besalú and coauthors derived the upper bound of the supremum norm and Hölder norm of the solution x in (2.3) without y. By a slight generalization of these results in [1,3], it is easy to obtain the following lemma.

Lemma 2.6 *(cf. [3, Theorem 4.1] and [1, Proposition 5.2]). Assume that $f \in C_b^1(\mathbb{R}^m \times \mathbb{R}^n;\mathbb{R}^m)$, $\sigma \in C_b^3(\mathbb{R}^m;\mathbb{R}^m \otimes \mathbb{R}^d)$. Then, we have the following estimate for the solution $(x,\omega,v) \in M_{m,d}^\beta$ of (2.3):*

$$\|x\|_\infty + \|x\|_\beta + \|v\|_{2\beta,\Delta} \le K_{\beta,T,|x_0|,f,\sigma}\Lambda_\omega^{\diamond(\beta)},$$

where $\diamond(\beta)$ denotes a certain positive constant which depends only on β and $K_{\beta,T,|x_0|,f,\sigma} > 0$ is a constant depending only on $\beta, T, |x_0|, \|f\|_\infty, \|\sigma\|_\infty, \|\nabla\sigma\|_\infty$ and $\|\nabla^2\sigma\|_\infty$ and is independent of y.

Proof. The proof of this result can be found in [1,3], although the paper [3] did not deal with any drift term f. However, such a term can be handled easily since the boundedness of f. Thus, by [3, Theorem 4.1 (i)] and Proposition 2.4 and Proposition 2.5, it is not difficulty to see that

$$\|x\|_\infty \le |x_0| + T(K\rho_{f,\sigma}\Lambda_\omega)^{\frac{1}{\beta}} + 1,$$

where $\rho_{f,\sigma} := \|f\|_\infty T^{1-\beta} + \|\sigma\|_\infty + \|\nabla\sigma\|_\infty + \|\nabla^2\sigma\|$ and K is a universal constant depending on β.

By [3, Step 2, pp. 251-252] or [1, Proposition 5.2], we can find a bound for the Hölder norms of $\|x\|_\beta$ and $\|v\|_{2\beta,\Delta}$, i.e. there exists a constant $K_{\beta,T,|x_0|,f,\sigma}$ depending only on $\beta, T, |x_0|, \|f\|_\infty, \|\sigma\|_\infty, \|\nabla\sigma\|_\infty$ and $\|\nabla^2\sigma\|_\infty$ such that

$$\|x\|_\beta + \|v\|_{2\beta,\Delta} \le K_{\beta,T,|x_0|,f,\sigma}\Lambda_\omega^{\diamond(\beta)}.$$

Thus, the statement holds. $\square$

In order to give some estimates involving a Lipschitz function σ, we need to introduce some notation:

$$\begin{aligned}
G^1_{\beta,[a,b]}(\sigma,x,\tilde{x},\omega,v) &= K[(\|\nabla^2\sigma\|_\infty + \|\nabla^2\sigma\|_\lambda(\|x\|^\lambda_{\beta,[a,b]} + \|\tilde{x}\|^\lambda_{\beta,[a,b]})(b-a)^{\lambda\beta}) \\
&\quad \times(\Phi_{\beta,[a,b]}(x,\omega,v) + \|\omega\|_{\beta,[a,b]}\|\tilde{x}\|_{\beta,[a,b]}) + \|\omega\|_{\beta,[a,b]}\|\nabla\sigma\|_\infty], \\
G^2_{\beta,[a,b]}(\sigma,x,\tilde{x},\omega,v) &= K[\|\nabla^2\sigma\|_\infty(\Phi_{\beta,[a,b]}(x,\omega,v) + \|\omega\|_{\beta,[a,b]}\|\tilde{x}\|_{\beta,[a,b]})(b-a)^\beta \\
&\quad +\|\omega\|_{\beta,[a,b]}\|\nabla\sigma\|_\infty], \\
G^3_{\beta,[a,b]}(\sigma,\tilde{x}) &= K[\|\nabla\sigma\|_\infty + \|\nabla^2\sigma\|_\infty\|\tilde{x}\|_{\beta,[a,b]}(b-a)^\beta].
\end{aligned}$$

Proposition 2.7 *(cf. [15, Proposition 6.4]). Let* $(x, \omega, v), (\tilde{x}, \omega, \tilde{v}) \in M^{\beta}_{m,d}$*. Assume that* $\sigma \in C_b^{2,\lambda}(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^d)$ *with* $(2+\lambda)\beta > 1$*. Then, for any* $(a, b) \in \Delta$*, we have*

$$\left| \int_a^b (\sigma(x_r) - \sigma(\tilde{x}_r)) \mathrm{d}\omega_r \right| \leq G^1_{\beta,[a,b]}(\sigma, x, \tilde{x}, \omega, v)(b-a)^{2\beta} \|x - \tilde{x}\|_{\infty,[a,b]}$$
$$+ G^2_{\beta,[a,b]}(\sigma, x, \tilde{x}, \omega, v)(b-a)^{2\beta} \|x - \tilde{x}\|_{\beta,[a,b]}$$
$$+ G^3_{\beta,[a,b]}(\sigma, \tilde{x})(b-a)^{2\beta} \|v - \tilde{v}\|_{2\beta,\Delta_{a,b}}.$$

Let us introduce more useful notations:

$$\begin{aligned}
G^4_{\beta,[a,b]}(\sigma, x, \tilde{x}, \omega, v, \omega^2) = K\{&\|\nabla\sigma\|_\infty \Phi_{\beta,[a,b]}(\omega, \omega, \omega^2) \\
&+ [\|\nabla^2\sigma\|_\infty + \|\nabla^2\sigma\|_\lambda (\|x\|^\lambda_{\beta,[a,b]} + \|\tilde{x}\|^\lambda_{\beta,[a,b]})(b-a)^{\lambda\beta}) \\
&\times (\Phi_{\beta,[a,b]}(x, \omega, v, \omega^2) + \|\tilde{x}\|_{\beta,[a,b]} \Phi_{\beta,[a,b]}(\omega, \omega, \omega^2)]\}, \\
G^5_{\beta,[a,b]}(\sigma, x, \tilde{x}, \omega, v, \omega^2) = K[&(\|\nabla\sigma\|_\infty + \|\nabla^2\sigma\|_\infty \|\tilde{x}\|_{\beta,[a,b]}(b-a)^\beta)\Phi_{\beta,[a,b]}(\omega, \omega, \omega^2) \\
&+ \|\nabla^2\sigma\|_\infty \Phi_{\beta,[a,b]}(x, \omega, v, \omega^2)(b-a)^\beta], \\
G^6_{\beta,[a,b]}(\sigma, \tilde{x}, \omega) = K &G^3_{\beta,[a,b]}(\sigma, \tilde{x}) \|\omega\|_{\beta,[a,b]}.
\end{aligned}$$

From the previous results it is possible to prove the following proposition.

Proposition 2.8 *(cf. [1, Proposition 4.9]). Let* $(x, \omega, v), (\tilde{x}, \omega, \tilde{v}) \in M^{\beta}_{m,d}$ *and* $(\omega, \omega^2) \in \mathscr{C}^\beta([0, T]; \mathbb{R}^d)$*. Assume that* $\sigma \in C_b^{2,\lambda}(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^d)$ *with* $(2+\lambda)\beta > 1$*. Then, for any* $(a, b) \in \Delta$*, we have*

$$\left| \int_a^b (\sigma(x_r) - \sigma(\tilde{x}_r)) \mathrm{d}\omega^2_{\cdot,b}(r) \right| \leq G^4_{\beta,[a,b]}(\sigma, x, \tilde{x}, \omega, v, \omega^2)(b-a)^{3\beta} \|x - \tilde{x}\|_{\infty,[a,b]}$$
$$+ G^5_{\beta,[a,b]}(\sigma, x, \tilde{x}, \omega, v, \omega^2)(b-a)^{3\beta} \|x - \tilde{x}\|_{\beta,[a,b]}$$
$$+ G^6_{\beta,[a,b]}(\sigma, \tilde{x}, \omega)(b-a)^{3\beta} \|v - \tilde{v}\|_{2\beta,\Delta_{a,b}}.$$

2.2. *Rough paths theory with approximation by Riemann sums*

In this subsection, we write a β-Hölder rough path by (X, X^2) instead of (ω, ω^2) following the notations by Friz and Hairer [9].

Definition 2.9. We say that a pair $(x, x^\dagger)$ is a controlled path with respect to (X, X^2) if the following decomposition holds

$$x_t - x_s = x^\dagger_s (X_t - X_s) + r_{st}, \quad (s, t) \in \Delta,$$

for certain $x^\dagger \in C^\beta([0,T];\mathbb{R}^m \otimes \mathbb{R}^d)$ and $r \in C_2^{2\beta}(\Delta;\mathbb{R}^m)$ where $x^\dagger$ is the Gubinelli derivative of x. The totality of such $(x, x^\dagger)$ is denoted by $\mathcal{D}_X^{2\beta}([0,T];\mathbb{R}^m)$, see e.g. [9, Definition 4.6]. We will omit the value space for simplicity of presentation, i.e. $\mathcal{D}_X^{2\beta}$ for shortness.

When $(x, x^\dagger) \in \mathcal{D}_X^{2\beta}$, then the rough integral of x against X can be defined as

$$\int_s^t x_r^i \mathrm{d}X_r^j = \lim_{|\mathcal{P}|\to 0} \sum_{t_k \in \mathcal{P}} \Big(x_{t_k}^i (X_{t_{k+1}}^j - X_{t_k}^j) + \sum_{\ell=1}^d (x_{t_k}^\dagger)^{i,\ell} X_{t_k t_{k+1}}^{2,\ell j} \Big), \tag{2.9}$$

where $i = 1, \dots m$, $j = 1, \dots, d$ and the limit is taken over all partitions $\mathcal{P} = \{s = t_{-1} = t_0 < t_1 < \dots < t_n = t_{n+1} = t\}$ such that $|\mathcal{P}| = \sup_{t_k \in \mathcal{P}} |t_k - t_{k-1}| \to 0$. It is known that $(\int_0^\cdot x_r^i \mathrm{d}X_r^j, x^i \mathbf{e}_j) \in \mathcal{D}_X^{2\beta}$, where $\{\mathbf{e}_j\}$ is the canonical basis of $\mathbb{R}^d$.

Proposition 2.10 *(cf. e.g. [9, Lemma 7.3]). Let $(x, x^\dagger) \in \mathcal{D}_X^{2\beta}$ and $\varphi \in C_b^2(\mathbb{R}^m;\mathbb{R}^n)$. Then we can define a controlled path $(\varphi(x), D\varphi(x)x^\dagger) \in \mathcal{D}_X^{2\beta}$, that is, $\varphi(x)$ is controlled by X if we take $\varphi(x)^\dagger = D\varphi(x)x^\dagger$ as a Gubinelli derivative of $\varphi(x)$, i.e.*

$$\varphi(x_t) - \varphi(x_s) - \varphi(x)_s^\dagger (X_t - X_s) \in C_2^{2\beta}(\Delta;\mathbb{R}^n).$$

Using appropriate estimates for the integrals, the solution to the following RDE driven by (X, X^2) in the sense of controlled paths theory:

$$x_t = x_0 + \int_0^t f(x_s)\mathrm{d}s + \int_0^t \sigma(x_s)\mathrm{d}X_s, \tag{2.10}$$

with $f : \mathbb{R}^m \to \mathbb{R}^m, \sigma : \mathbb{R}^m \to \mathbb{R}^m \otimes \mathbb{R}^d$ is obtained via a fixed point argument.

Lemma 2.11 *(cf. e.g. [9, Section 8]). Suppose that $(X, X^2) \in \mathscr{C}^\beta([0,T];\mathbb{R}^d)$ and let $f \in C_b^1(\mathbb{R}^m, \mathbb{R}^m)$ and $\sigma \in C_b^3(\mathbb{R}^m, \mathbb{R}^m \otimes \mathbb{R}^d)$. Then the equation (2.10) possesses a unique global solution $(x, \sigma(x)) \in \mathcal{D}_X^{2\beta}([0,T], \mathbb{R}^m)$. Here, both sides are understood as elements of $\mathcal{D}_X^{2\beta}([0,T], \mathbb{R}^m)$.*

As a consequence of this Proposition 2.10 and Lemma 2.11, we have

$$\int_0^t \sigma_j(x_r)\mathrm{d}X_r^j = \lim_{|\mathcal{P}|\to 0} \sum_{t_k \in \mathcal{P}} \Big(\sigma_j(x_{t_k})(X_{t_{k+1}}^j - X_{t_k}^j) + \sum_{\ell=1}^d \mathcal{D}_\sigma^{(\ell)} \sigma_j(x_{t_k}) X_{t_k t_{k+1}}^{2,\ell j} \Big), \tag{2.11}$$

with $\sigma_j : \mathbb{R}^m \to \mathbb{R}^m$, $j = 1, 2 \dots d$, where the differential operators $\mathcal{D}_\sigma^{(\ell)} = \sum_{l=1}^m \sigma_{l,\ell}(\cdot)\partial_{x_l}$. It is known that $(x_0 + \int_0^\cdot f(x_s)\mathrm{d}s + \int_0^\cdot \sigma_j(x_r)\mathrm{d}X_r^j, \sigma_j(x_\cdot)) \in \mathcal{D}_X^{2\beta}([0,T], \mathbb{R}^m)$.

Remark 2.12. Let $\sigma \in C_b^3(\mathbb{R}^m;\mathbb{R}^m \otimes \mathbb{R}^d)$.

- For $(x, \omega, v) \in M^{\beta}_{m,d}$, the integral $\int_0^T \sigma(x_s)\mathrm{d}\omega_s$ coincides with the integral defined using the $\frac{1}{\beta}$-variation norm (see Ito [16, Theorem 2.5 and Remark 2.6]). This implies that $\int_0^T \sigma(x_s)\mathrm{d}\omega_s$ in Definition 2.2 can be given by the limit of Riemann sums of the same form of given in (2.11).
- The first level of Hu-Nualart type (Definition 2.2) and Gubinelli type (Definition 2.11) unique global solution are known to coincide.

Using the flow method, the solution to the RDE with unbounded drift term is obtained by Riedel and Scheutzow [31].

Lemma 2.13 *(cf. e.g. [31, Theorem 3.1]). Suppose that $(X, X^2) \in \mathscr{C}^{\beta}_g([0, T]; \mathbb{R}^d)$ and assume that f is a locally Lipschitz continuous vector field with at most linear growth on $\mathbb{R}^m$ and $\sigma \in C^4_b(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^d)$. Then a unique global solution exists for any initial value x_0.*

3. RDEs driven by mixed fractional Brownian rough path

We mainly use rough path theory recalled in Section 2 to prove that (1.1) possesses a unique global solution. To do that, we recall the following lemma.

Lemma 3.1 *(cf. [24, Lemma 2]). Suppose $(S(t), \mathcal{F}_t)_{t\in[0,T]}$ is a stochastic process with β-Hölder trajectories for all $\beta \in (\frac{1}{3}, \frac{1}{2})$, such that $\mathbb{E}[\|S\|^p_\beta] < \infty$ for all $p \geq 1$. Then, for all $\beta' \in (0, \beta)$, there exist a modification of $(s, t) \to \int_s^t (S(u) - S(s))\mathrm{d}^{\mathrm{I}} W_u$ and an almost surely finite random variable $C_{T,\beta'}$ such that*

$$\left| \int_s^t (S(u) - S(s))\mathrm{d}^{\mathrm{I}} W_u \right| \leq C_{T,\beta'} |t - s|^{\frac{1}{2}+\beta'}, \quad (s, t) \in \Delta.$$

Now, we are ready to prove Proposition 1.1.

Proof of Proposition 1.1. We understand (1.1) as following RDE

$$u_t = u_0 + \int_0^t a(u_s)\mathrm{d}s + \int_0^t (b, c)(u_s)\mathrm{d}\mathbf{Z}_s,$$

where $\mathbf{Z} = (Z, Z^2)$ is a joint, step-2 rough path lift between the Bm $\mathbf{W}$ and fBm $\mathbf{B}$ (which will be defined below). Here (b, c) is the $m \times (d + d')$ block matrix. Set $Z_{st} \triangleq Z_t - Z_s$ and denote

$$Z_{st} = (B_{st}, W_{st})^{\mathrm{T}}, \quad Z^2_{st} = \begin{pmatrix} B^2_{st} & I[B, W]_{st} \\ I[W, B]_{st} & W^2_{st} \end{pmatrix},$$

where $\mathbf{B} = (B, B^2)$ is a canonical geometric rough path (see Friz and Hairer [9, Section 10.3] for example) associated to fBm, $\mathbf{W} = (W, W^2)$ is a geometric rough path in Stratonovich sense (see

Friz and Hairer [9, Section 3] for example) associated to Bm, and we define for every $(s,t) \in \Delta$,

$$I[B,W]_{st} \triangleq \int_s^t B_{su} \otimes \mathrm{d}^{\mathrm{I}} W_u,$$

$$I[W,B]_{st} \triangleq W_{st} \otimes B_{st} - \int_s^t \mathrm{d}^{\mathrm{I}} W_u \otimes B_{su}.$$

Firstly, we need to prove the joint, step-2 rough path lift $\mathbf{Z}$ satisfies (2.2). For $Z^{2,ij}_{st}$, $i,j \in \{1,\dots,d\}$ and $i,j \in \{d+1,\dots,d+d'\}$, it is easy to know that Z^2 satisfy the Chen's relation. It remains to demonstrate that when $i \in \{1,\dots,d\}$, $j \in \{d+1,\dots,d+d'\}$ and $j \in \{1,\dots,d\}, i \in \{d+1,\dots,d+d'\}$, whether we can obtain same relation. Let us study

$$Z^{2,ij}_{st} = \int_s^t B^i_{su} \mathrm{d}^{\mathrm{I}} W^j_u, \quad i \in \{1,\dots,d\}, j \in \{d+1,\dots,d+d'\},$$

then, if $i \in \{1,\dots,d\}$, $j \in \{d+1,\dots,d+d'\}$, we have for $s \le u \le t$,

$$\begin{aligned} Z^{2,ij}_{st} - Z^{2,ij}_{su} - Z^{2,ij}_{ut} &= \int_s^t B^i_{sr} \mathrm{d}^{\mathrm{I}} W^j_r - \int_s^u B^i_{sr} \mathrm{d}^{\mathrm{I}} W^j_r - \int_u^t B^i_{ur} \mathrm{d}^{\mathrm{I}} W^j_r \\ &= \int_u^t B^i_{su} \mathrm{d}^{\mathrm{I}} W^j_r \\ &= Z^i_{su} Z^j_{ut}. \end{aligned} \tag{3.1}$$

By (3.1), it is easy to obtain $Z^{2,ij}_{st} - Z^{2,ij}_{su} - Z^{2,ij}_{ut} = Z^i_{su} Z^j_{ut}$ holds for $j \in \{1,\dots,d\}, i \in \{d+1,\dots,d+d'\}$. Thus, the joint, step-2 rough path lift $\mathbf{Z}$ satisfies (2.2).

Then, there remains the analytic condition to be checked. By Lemma 3.1, it follows that almost surely $(Z, Z^2) \in \mathscr{C}^\beta([0,T];\mathbb{R}^{d+d'})$ for any $\beta \in (\frac{1}{3}, H)$. It is easy to check $Z^{2,ij}_{st} + Z^{2,ji}_{st} = Z^i_{st} Z^j_{st}$. So, $(Z, Z^2) \in \mathscr{C}^\beta_g([0,T];\mathbb{R}^{d+d'})$ for any $\beta \in (\frac{1}{3}, H)$.

Finally, according to Lemma 2.13, because a is a locally Lipschitz continuous vector field with at most linear growth on $\mathbb{R}^e$ and $b_i, c_j \in C^4_b(\mathbb{R}^e, \mathbb{R}^e)$ $(1 \le i \le d, 1 \le j \le d')$. Then, (1.1) possesses a unique global solution. By Lemma 2.11, if $a \in C^1_b(\mathbb{R}^e, \mathbb{R}^e)$ and $b_i, c_j \in C^3_b(\mathbb{R}^e, \mathbb{R}^e)$ $(1 \le i \le d, 1 \le j \le d')$. Then, (1.1) possesses a unique global solution. □

Through a similar argument as in the proof of Proposition 1.1, we prove that (1.3) possesses a unique global solution.

Theorem 3.2. *Let $\beta \in (\frac{1}{3}, H)$ and assume the coefficients of (1.3) satisfy* (H1) *and* (H2). *Then, (1.3) possesses a unique global solution.*

Proof. This is just the special case of Proposition 1.1. Let

$$u_t^\varepsilon := \begin{pmatrix} X_t^\varepsilon \\ Y_t^\varepsilon \end{pmatrix}, \quad F(\xi,\phi) := \begin{pmatrix} f(\xi,\phi) \\ \frac{1}{\varepsilon} g(\xi,\phi) \end{pmatrix} \quad \text{and} \quad V(\xi,\phi) := \begin{pmatrix} \sigma(\xi) & O \\ O & \frac{1}{\sqrt{\varepsilon}} h(\xi,\phi) \end{pmatrix}, \tag{3.2}$$

where $F : \mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$ and $V = (V_1, \dots, V_{d+d'})$ is $(m+n) \times (d+d')$ matrix-valued. Then, set $(b,c) = V$ and by (H1) and (H2), (1.3) possesses a unique global solution. $\square$

Moreover, by (1.5), it is easy to know $\bar{f}$ is also a Lipschitz continuous vector field with at most linear growth, thus (1.4) possesses a unique global solution.

Now, we study the relation between the fast component of RDE (1.3) and an Itô SDE. Note that $\mathbb{E}^B, \mathbb{E}^W$ are the expectation with respect to B and W, respectively, so that $\mathbb{E} = \mathbb{E}^B \times \mathbb{E}^W$.

Theorem 3.3. *The first level path of the fast component of RDE (1.3) is the following Itô SDE*

$$Y_t^\varepsilon = Y_0 + \frac{1}{\varepsilon} \int_0^t \tilde{g}(X_s^\varepsilon, Y_s^\varepsilon) ds + \frac{1}{\sqrt{\varepsilon}} \int_0^t h(X_s^\varepsilon, Y_s^\varepsilon) d^{\mathrm{I}} W_s. \tag{3.3}$$

Here, X^ε is the first level path of the slow component of RDE (1.3) and $\tilde{g}$ has been defined in (1.6).

Proof. By (2.11), we rewrite the rough path lift terms of the right side of (1.3) as

$$\begin{aligned} &\sum_{l=1}^{m+n} \sum_{j=1}^{d+d'} \int_0^t V_{l,j}(u_s^\varepsilon) dZ_s^j \\ &= \lim_{|\mathcal{P}|\to 0} \sum_{t_k \in \mathcal{P}} \sum_{l=1}^{m+n} \sum_{j=1}^{d+d'} \left\{ V_{l,j}(u_{t_k}^\varepsilon) Z_{t_k t_{k+1}}^j + \sum_{i=1}^{d+d'} \mathcal{D}_V^{(i)} V_{l,j}(u_{t_k}^\varepsilon) Z_{t_k t_{k+1}}^{2,ij} \right\}, \end{aligned} \tag{3.4}$$

where $\mathcal{D}_V^{(i)} = \sum_{l=1}^{m+n} V_{l,i}(\cdot) \partial_{u_l}, i = 1, \dots, d+d'$.

To prove the theorem, it is sufficient to compute the fast component Y^ε of (1.3). According to (3.4), taking $(m+1) \le l \le (m+n)$, we have

$$\begin{aligned} &\lim_{|\mathcal{P}|\to 0} \sum_{t_k \in \mathcal{P}} \sum_{l=m+1}^{m+n} \sum_{j=1}^{d+d'} \left\{ V_{l,j}(u_{t_k}^\varepsilon) Z_{t_k t_{k+1}}^j + \sum_{i=1}^{d+d'} \mathcal{D}_V^{(i)} V_{l,j}(u_{t_k}^\varepsilon) Z_{t_k t_{k+1}}^{2,ij} \right\} \\ &= \lim_{|\mathcal{P}|\to 0} \sum_{t_k \in \mathcal{P}} \left\{ \sum_{\bar{l}=1}^{n} \sum_{\bar{j}=1}^{d'} h_{\bar{l},\bar{j}}(X_{t_k}^\varepsilon, Y_{t_k}^\varepsilon) W_{t_k t_{k+1}}^{\bar{j}} + \sum_{l=m+1}^{m+n} \sum_{j=d+1}^{d+d'} \sum_{i=1}^{d} \mathcal{D}_V^{(i)} V_{l,j}(u_{t_k}^\varepsilon) Z_{t_k t_{k+1}}^{2,ij} \right. \\ &\quad \left. + \sum_{l=m+1}^{m+n} \sum_{j=d+1}^{d+d'} \sum_{i=d+1}^{d+d'} \mathcal{D}_V^{(i)} V_{l,j}(u_{t_k}^\varepsilon) Z_{t_k t_{k+1}}^{2,ij} \right\} =: \lim_{|\mathcal{P}|\to 0} \sum_{i=1}^{3} \mathbf{I}_i. \end{aligned}$$

We will prove that $\lim_{|\mathcal{P}|\to 0} \mathbf{I}_i$ exists for $i = 1, 2, 3$ in the sense of limit in probability.

For the term $\mathbf{I}_1$, it is easy to obtain

$$\lim_{|\mathcal{P}|\to 0} \sum_{t_k\in\mathcal{P}} \sum_{\bar{l}=1}^{n} \sum_{\bar{j}=1}^{d'} h_{\bar{l},\bar{j}}(X^\varepsilon_{t_k}, Y^\varepsilon_{t_k}) W^{\bar{j}}_{t_k t_{k+1}} = \sum_{\bar{l}=1}^{n} \sum_{\bar{j}=1}^{d'} \int_0^t h_{\bar{l},\bar{j}}(X^\varepsilon_s, Y^\varepsilon_s) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_s \quad \text{in } L^2,$$

by the definition of the Itô integral because $h_{\bar{l},\bar{j}}(X^\varepsilon_s, Y^\varepsilon_s)$, $s \in [0, T]$, is continuous, bounded and adapted.

For the second term $\mathbf{I}_2$, set $\mathcal{D}^{(\hat{i})}_\sigma = \sum_{\hat{l}=1}^{m} \sigma_{\hat{l},\hat{i}}(\cdot)\partial_{x_{\hat{l}}}$, there exists a constant $K > 0$, one has

$$\begin{aligned}
\mathbb{E}^W[\mathbf{I}_2^2] &= \mathbb{E}^W\Big[\Big(\sum_{t_k\in\mathcal{P}} \sum_{l=m+1}^{m+n} \sum_{j=d+1}^{d+d'} \sum_{i=1}^{d} \mathcal{D}^{(i)}_V V_{l,j}(u^\varepsilon_{t_k}) Z^{2,ij}_{t_k t_{k+1}}\Big)^2\Big] \\
&= \mathbb{E}^W\Big[\Big(\sum_{t_k\in\mathcal{P}} \sum_{\bar{l}=1}^{n} \sum_{\bar{j}=1}^{d'} \sum_{\hat{i}=1}^{d} \mathcal{D}^{(\hat{i})}_\sigma h_{\bar{l},\bar{j}}(X^\varepsilon_{t_k}, Y^\varepsilon_{t_k}) I[B, W]^{\hat{i},\bar{j}}_{t_k t_{k+1}}\Big)^2\Big] \\
&\le K \sum_{\bar{l}=1}^{n} \sum_{\bar{j}=1}^{d'} \sum_{\hat{i}=1}^{d} \mathbb{E}^W\Big[\Big(\sum_{t_k\in\mathcal{P}} \mathcal{D}^{(\hat{i})}_\sigma h_{\bar{l},\bar{j}}(X^\varepsilon_{t_k}, Y^\varepsilon_{t_k}) \int_{t_k}^{t_{k+1}} (B^{\hat{i}}_r - B^{\hat{i}}_{t_k}) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_r\Big)^2\Big].
\end{aligned}$$

To prove this, put $A_{\hat{i},\bar{l},\bar{j}}(t) := \mathcal{D}^{(\hat{i})}_\sigma h_{\bar{l},\bar{j}}(X^\varepsilon_t, Y^\varepsilon_t)$, $A_{\hat{i},\bar{l},\bar{j}}(k) := A_{\hat{i},\bar{l},\bar{j}}(t_k)$ and consider

$$\begin{aligned}
&\mathbb{E}^W\Big[\Big(\sum_k A_{\hat{i},\bar{l},\bar{j}}(k) \int_{t_k}^{t_{k+1}} (B^{\hat{i}}_r - B^{\hat{i}}_{t_k}) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_r\Big)^2\Big] \\
&\quad = \mathbb{E}^W\Big[\sum_{k,k'} A_{\hat{i},\bar{l},\bar{j}}(k) A_{\hat{i},\bar{l},\bar{j}}(k')\Big(\int_{t_k}^{t_{k+1}} (B^{\hat{i}}_r - B^{\hat{i}}_{t_k}) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_r\Big)\Big(\int_{t_{k'}}^{t_{k'+1}} (B^{\hat{i}}_r - B^{\hat{i}}_{t_{k'}}) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_r\Big)\Big].
\end{aligned}$$

If $k < k'$ then $A_{\hat{i},\bar{l},\bar{j}}(k) A_{\hat{i},\bar{l},\bar{j}}(k')\big(\int_{t_k}^{t_{k+1}} (B^{\hat{i}}_r - B^{\hat{i}}_{t_k}) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_r\big)$ and $\big(\int_{t_{k'}}^{t_{k'}+1} (B^{\hat{i}}_r - B^{\hat{i}}_{t_{k'}}) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_r\big)$ are independent so the terms vanish in this case, and similarly if $k > k'$. So we are left with

$$\mathbb{E}^W[\mathbf{I}_2^2] \le K \sum_{\bar{l}=1}^{n} \sum_{\bar{j}=1}^{d'} \sum_{\hat{i}=1}^{d} \sum_{t_k\in\mathcal{P}} \mathbb{E}^W\Big[\Big(\mathcal{D}^{(\hat{i})}_\sigma h_{\bar{l},\bar{j}}(X^\varepsilon_{t_k}, Y^\varepsilon_{t_k}) \int_{t_k}^{t_{k+1}} (B^{\hat{i}}_r - B^{\hat{i}}_{t_k}) \mathrm{d}^{\mathrm{I}} W^{\bar{j}}_r\Big)^2\Big].$$

Note that $\mathbb{E}^B\big[\|B^{\hat{i}}\|^2_\beta\big] < \infty$, thus, one has $\mathbb{E}[\hat{\mathbf{I}}_2^2] = \mathbb{E}^B[\mathbb{E}^W[\mathbf{I}_2^2]] \to 0$ as $|\mathcal{P}| \to 0$.

To proceed, for the third term $\mathbf{I}_3$, one has

$$\lim_{|\mathcal{P}|\to 0} \mathbf{I}_3 = \lim_{|\mathcal{P}|\to 0} \sum_{t_k\in\mathcal{P}} \sum_{\bar{l}=1}^{n} \sum_{\bar{i}=1}^{d'} \sum_{\bar{j}=1}^{d'} \mathcal{D}^{(\bar{i})}_h h_{\bar{l},\bar{j}}(X^\varepsilon_{t_k}, Y^\varepsilon_{t_k}) \big(W^2_{t_k t_{k+1}}\big)^{\bar{i},\bar{j}},$$

where $\mathcal{D}_h^{(\bar{i})} = \sum_{\bar{l}=1}^{n} h_{\bar{l},\bar{i}}(\cdot,\cdot)\partial_{y_{\bar{l}}}$. Following [24, Theorem 4], for $\bar{i} = \bar{j}$, we have

$$\lim_{|\mathcal{P}|\to 0} \sum_{t_k\in\mathcal{P}} \sum_{\bar{l}=1}^{n} \sum_{\bar{i}=1}^{d'} \mathcal{D}_h^{(\bar{i})} h_{\bar{l},\bar{i}}(X_{t_k}^{\varepsilon}, Y_{t_k}^{\varepsilon})\big(W_{t_k t_{k+1}}^2\big)^{\bar{i},\bar{i}} = \frac{1}{2} \sum_{\bar{l}=1}^{n} \sum_{\bar{i}=1}^{d'} \int_0^t \mathcal{D}_h^{(\bar{i})} h_{\bar{l},\bar{i}}(X_s^{\varepsilon}, Y_s^{\varepsilon}) \mathrm{d}s, \quad \text{a.s.}$$

Then, by [24, Theorem 4] again, for $\bar{i} \neq \bar{j}$, we obtain

$$\lim_{|\mathcal{P}|\to 0} \mathbb{E}\Bigg[\bigg(\sum_{t_k\in\mathcal{P}} \sum_{\bar{l}=1}^{n} \sum_{\bar{i}=1}^{d'} \sum_{\bar{j}=1}^{d'} \mathcal{D}_h^{(\bar{i})} h_{\bar{l},\bar{j}}(X_{t_k}^{\varepsilon}, Y_{t_k}^{\varepsilon})\big(W_{t_k t_{k+1}}^2\big)^{\bar{i},\bar{j}}\bigg)^2\Bigg] = 0 \quad \text{in } L^2.$$

Thus, we have shown (3.3). $\square$

4. Averaging principle

In this section, we combine the pathwise approach via fractional calculus and rough path theory to estimate the slow component X^{ε} and fast component Y^{ε} of RDE (1.3), respectively. Now, let us study the slow component of RDE (1.3) using the pathwise approach via fractional calculus. By (3.2), it is not difficult to show that there exists a triplet $(u^{\varepsilon}, Z, \tilde{v}^{\varepsilon}) \in M_{m+n,d+d'}^{\beta}$ (this section follows the notations proposed by Hu and Nualart [15]). Note that the slow component X^{ε} is the solution of (2.3) with (y, ω) replaced by (Y^{ε}, B). In particular, its "v-component" is of the following form

$$v_{st}^{\varepsilon} = \int_s^t \int_s^r f(X_q^{\varepsilon}, Y_q^{\varepsilon}) \mathrm{d}q \otimes \mathrm{d}B_r - \int_s^t \sigma(X_r^{\varepsilon}) \mathrm{d}B_{\cdot,t}^2(r), \tag{4.1}$$

where the last integral is defined based on fractional calculus theory (see Definition 2.2) and (4.1) is well defined under the conditions (H2) and (H3). The stochastic integral of slow component of (1.3) is a pathwise integral which depends on B and B^2.

Following the discretization techniques inspired by Khasminskii in [17], we introduce an auxiliary process $(\hat{X}^{\varepsilon}, \hat{Y}^{\varepsilon})$ and divide $[0, T]$ into intervals depending of size $\delta < 1$, where δ is a fixed positive number depending on ε which will be chosen later. Then, we construct $\hat{Y}^{\varepsilon}$ with initial value $\hat{Y}_0^{\varepsilon} = Y_0$,

$$\hat{Y}_t^{\varepsilon} = Y_0 + \frac{1}{\varepsilon} \int_0^t \tilde{g}(X_{s(\delta)}^{\varepsilon}, \hat{Y}_s^{\varepsilon}) \mathrm{d}s + \frac{1}{\sqrt{\varepsilon}} \int_0^t h(X_{s(\delta)}^{\varepsilon}, \hat{Y}_s^{\varepsilon}) \mathrm{d}^{\mathrm{I}} W_s, \tag{4.2}$$

where $s(\delta) = \lfloor s\delta^{-1} \rfloor \delta$ is the nearest breakpoint preceding s. Also, we define the process $\hat{X}^{\varepsilon}$ with initial value $\hat{X}_0^{\varepsilon} = X_0$, by

$$\hat{X}_t^{\varepsilon} = X_0 + \int_0^t f(X_{s(\delta)}^{\varepsilon}, \hat{Y}_s^{\varepsilon}) \mathrm{d}s + \int_0^t \sigma(X_s^{\varepsilon}) \mathrm{d}B_s, \tag{4.3}$$

in the rough path sense. Similarly, denote the second component of (4.3) and (1.4) as $\hat{v}^\varepsilon$ and $\bar{v}$, i.e. $(\hat{X}^\varepsilon, B, \hat{v}^\varepsilon), (\bar{X}, B, \bar{v}) \in M^\beta_{m,d}$, respectively.

4.1. Some estimates on the solutions X^ε, $\hat{X}^\varepsilon$, $\bar{X}$, Y^ε and $\hat{Y}^\varepsilon$

From now on, $\diamond(\beta)$ is a certain positive constant and may vary from line to line. Then, by Lemma 2.6, it is easy to obtain the following lemmas.

Lemma 4.1. *Assume that f, σ satisfy* (H1)-(H3). *Then, we have the following estimate:*

$$\|X^\varepsilon\|_\infty + \|X^\varepsilon\|_\beta + \|v^\varepsilon\|_{2\beta,\Delta} \le K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)},$$

almost surely, where $K_{\beta,T,|X_0|,f,\sigma} > 0$ is a constant depending only on $\beta, T, |X_0|, \|f\|_\infty, \|\sigma\|_\infty$, $\|\nabla\sigma\|_\infty$ and $\|\nabla^2\sigma\|_\infty$ and Λ_B has moments of all order.

Using similar techniques, the statements proposed in Lemma 4.1 also hold for $\hat{X}^\varepsilon$ and $\bar{X}$.

Lemma 4.2. *Assume that f, σ satisfy* (H1)-(H3). *Then, for $(t, t+\delta) \in \Delta$, we have the following estimate:*

$$\sup_{t\in[0,T]} |X^\varepsilon_{t+\delta} - X^\varepsilon_t| \le K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta,$$

almost surely.

Lemma 4.3. *Assume that f, σ satisfy* (H1)-(H3) *and let $(X^\varepsilon, B, v^\varepsilon), (\hat{X}^\varepsilon, B, \hat{v}^\varepsilon) \in M^\beta_{m,d}$ be as in (1.3) and (4.3), respectively. Then, we have the following estimates:*

$$\|X^\varepsilon - \hat{X}^\varepsilon\|_\infty \le K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + \Phi_\varepsilon,$$
$$\|X^\varepsilon - \hat{X}^\varepsilon\|_{\beta,[a,b]} \le K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + \Phi_\varepsilon,$$
$$\|v^\varepsilon - \hat{v}^\varepsilon\|_{2\beta,\Delta_{a,b}} \le \|B\|_\beta(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + \Phi_\varepsilon),$$

almost surely, where

$$\Phi_\varepsilon = \left\| \int_0^\cdot (f(X^\varepsilon_{r(\delta)}, \hat{Y}^\varepsilon_r) - f(X^\varepsilon_{r(\delta)}, Y^\varepsilon_r))\mathrm{d}r \right\|_\infty + \left\| \int_0^\cdot (f(X^\varepsilon_{r(\delta)}, \hat{Y}^\varepsilon_r) - f(X^\varepsilon_{r(\delta)}, Y^\varepsilon_r))\mathrm{d}r \right\|_\beta .$$

Proof. We start studying the supremum norm. By Lemma 4.2 and (H3), we obtain

$$\|X^\varepsilon - \hat{X}^\varepsilon\|_\infty = \left\| \int_0^\cdot (f(X^\varepsilon_s, Y^\varepsilon_s) - f(X^\varepsilon_{s(\delta)}, \hat{Y}^\varepsilon_s))\mathrm{d}s \right\|_\infty$$
$$\le \sup_{t\in[0,T]} \int_0^t |f(X^\varepsilon_s, Y^\varepsilon_s) - f(X^\varepsilon_{s(\delta)}, Y^\varepsilon_s)|\mathrm{d}s$$

$$+\left\| \int_0^{\cdot} (f(X^\varepsilon_{s(\delta)}, \hat{Y}^\varepsilon_s) - f(X^\varepsilon_{s(\delta)}, Y^\varepsilon_s))\mathrm{d}s \right\|_\infty$$
$$\leq K_{\beta,T,|X_0|,f,\sigma} \Lambda_B^{\diamond(\beta)} \delta^\beta + \Phi_\varepsilon, \quad \text{a.s.}$$

Now, we study the Hölder norm. Using Lemma 4.2 and by (H3) again, we obtain

$$\|X^\varepsilon - \hat{X}^\varepsilon\|_{\beta,[a,b]} = \sup_{(s,t)\in\Delta_{a,b}} \frac{|\int_s^t (f(X^\varepsilon_r, Y^\varepsilon_r) - f(X^\varepsilon_{r(\delta)}, \hat{Y}^\varepsilon_r))\mathrm{d}r|}{(t-s)^\beta}$$
$$\leq K_{\beta,T,|X_0|,f,\sigma} \Lambda_B^{\diamond(\beta)} (b-a)^{1-\beta} \delta^\beta + \Phi_\varepsilon, \quad \text{a.s.}$$

Now, we study the Hölder norm $\|v^\varepsilon - \hat{v}^\varepsilon\|_{2\beta,\Delta_{a,b}}$. By (4.1), Fubini's theorem and the argument proposed in [10, p. 2367], we have

$$\|v^\varepsilon - \hat{v}^\varepsilon\|_{2\beta,\Delta_{a,b}} = \sup_{(s,t)\in\Delta_{a,b}} \frac{|\int_s^t \int_s^r (f(X^\varepsilon_q, Y^\varepsilon_q) - f(X^\varepsilon_{q(\delta)}, \hat{Y}^\varepsilon_q))\mathrm{d}q \otimes \mathrm{d}B_r|}{(t-s)^{2\beta}}$$
$$\leq \|B\|_\beta (K_{\beta,T,|X_0|,f,\sigma} \Lambda_B^{\diamond(\beta)} (b-a)^{1-\beta} \delta^\beta + \Phi_\varepsilon), \quad \text{a.s.}$$

This completed the proof of Lemma 4.3. $\square$

Next, let us study $\|\hat{X}^\varepsilon - \bar{X}\|_\infty$.

Lemma 4.4. *Assume that* f, σ *satisfy* (H1)-(H3) *and let* $(\hat{X}^\varepsilon, B, \hat{v}^\varepsilon), (\bar{X}, B, \bar{v}) \in M^\beta_{m,d}$ *be as in (4.3) and (1.4), respectively. Then, we have the following estimates:*

$$\|\hat{X}^\varepsilon - \bar{X}\|_\infty \leq K_{\beta,T,|X_0|,f,\sigma} 2^{\Lambda_B^{\diamond(\beta)}} \Lambda_B^{\diamond(\beta)} (\delta^\beta + \mathbf{A}_1 + \mathbf{B}_1 + \Phi_\varepsilon),$$

almost surely, where

$$\mathbf{A}_1 = \left\| \int_0^{\cdot} (f(X^\varepsilon_{s(\delta)}, \hat{Y}^\varepsilon_s) - \bar{f}(X^\varepsilon_{s(\delta)}))\mathrm{d}s \right\|_\beta,$$
$$\mathbf{B}_1 = \sup_{(s,t)\in\Delta} \frac{|\int_s^t \int_s^r (f(X^\varepsilon_{q(\delta)}, \hat{Y}^\varepsilon_q) - \bar{f}(X^\varepsilon_{q(\delta)}))\mathrm{d}q \otimes \mathrm{d}B_r|}{(t-s)^{2\beta}}.$$

Proof. Fix a realization of (B, B^2), then everything in this proof is deterministic. Our first purpose is to estimate the Hölder norm $\|\hat{X}^\varepsilon - \bar{X}\|_{\beta,[a,b]}$. By (1.4) and (4.3), we have

$$\|\hat{X}^\varepsilon - \bar{X}\|_{\beta,[a,b]} \leq \left\| \int_0^{\cdot} (f(X^\varepsilon_{s(\delta)}, \hat{Y}^\varepsilon_s) - \bar{f}(X^\varepsilon_{s(\delta)}))\mathrm{d}s \right\|_\beta$$

$$
\begin{aligned}
&+\left\| \int_a^{\cdot} (\bar{f}(X^\varepsilon_{s(\delta)}) - \bar{f}(X^\varepsilon_s))\mathrm{d}s \right\|_{\beta,[a,b]} \\
&+\left\| \int_a^{\cdot} (\bar{f}(X^\varepsilon_s) - \bar{f}(\hat{X}^\varepsilon_s))\mathrm{d}s \right\|_{\beta,[a,b]} \\
&+\left\| \int_a^{\cdot} (\bar{f}(\hat{X}^\varepsilon_s) - \bar{f}(\bar{X}_s))\mathrm{d}s \right\|_{\beta,[a,b]} \\
&+\left\| \int_a^{\cdot} (\sigma(\hat{X}^\varepsilon_s) - \sigma(\bar{X}_s))\mathrm{d}B_s \right\|_{\beta,[a,b]} \\
&+\left\| \int_a^{\cdot} (\sigma(X^\varepsilon_s) - \sigma(\hat{X}^\varepsilon_s))\mathrm{d}B_s \right\|_{\beta,[a,b]} \\
=: &\sum_{i=1}^{6} \mathbf{A}_i.
\end{aligned}
$$

Let us study $\mathbf{A}_2$, $\mathbf{A}_3$ and $\mathbf{A}_4$. By Lemma 4.3, it is easy to obtain

$$
\begin{aligned}
\sum_{i=2}^{4} \mathbf{A}_i &\le K(b-a)^{1-\beta}\Big(\sup_{t\in[0,T]} |X^\varepsilon_t - X^\varepsilon_{t(\delta)}| + \|X^\varepsilon - \hat{X}^\varepsilon\|_\infty + \|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]}\Big) \\
&\le K(b-a)^{1-\beta}\big(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + \Phi_\varepsilon + \|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]}\big).
\end{aligned}
$$

Now we estimate $\mathbf{A}_5$ and $\mathbf{A}_6$. Since $\sigma \in C_b^3(\mathbb{R}^m; \mathbb{R}^m \otimes \mathbb{R}^d)$, taking $\lambda = 1$ in Proposition 2.7, we have

$$
\begin{aligned}
\mathbf{A}_5 \le\; & G^1_{\beta,[a,b]}(\sigma, \hat{X}^\varepsilon, \bar{X}, B, \hat{v}^\varepsilon)(b-a)^\beta \|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]} \\
&+ G^2_{\beta,[a,b]}(\sigma, \hat{X}^\varepsilon, \bar{X}, B, \hat{v}^\varepsilon)(b-a)^\beta \|\hat{X}^\varepsilon - \bar{X}\|_{\beta,[a,b]} \\
&+ G^3_{\beta,[a,b]}(\sigma, \bar{X})(b-a)^\beta \|\hat{v}^\varepsilon - \bar{v}\|_{2\beta,\Delta_{a,b}},
\end{aligned}
$$

and by Proposition 2.7, Lemma 4.2 and Lemma 4.3, we have

$$
\begin{aligned}
\mathbf{A}_6 \le\; & G^1_{\beta,[a,b]}(\sigma, X^\varepsilon, \hat{X}^\varepsilon, B, v^\varepsilon)(b-a)^\beta \|X^\varepsilon - \hat{X}^\varepsilon\|_{\infty,[a,b]} \\
&+ G^2_{\beta,[a,b]}(\sigma, X^\varepsilon, \hat{X}^\varepsilon, B, v^\varepsilon)(b-a)^\beta \|X^\varepsilon - \hat{X}^\varepsilon\|_{\beta,[a,b]} \\
&+ G^3_{\beta,[a,b]}(\sigma, \hat{X}^\varepsilon)(b-a)^\beta \|v^\varepsilon - \hat{v}^\varepsilon\|_{2\beta,\Delta_{a,b}} \\
\le\; & G^1_{\beta,[a,b]}(\sigma, X^\varepsilon, \hat{X}^\varepsilon, B, v^\varepsilon)(b-a)^\beta (K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + \Phi_\varepsilon) \\
&+ G^2_{\beta,[a,b]}(\sigma, X^\varepsilon, \hat{X}^\varepsilon, B, v^\varepsilon)(b-a)^\beta (K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + \Phi_\varepsilon) \\
&+ G^3_{\beta,[a,b]}(\sigma, \hat{X}^\varepsilon)\|B\|_\beta (b-a)^\beta (K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + \Phi_\varepsilon).
\end{aligned}
$$

Putting above estimations together, we obtain

$$\begin{aligned}\|\hat{X}^\varepsilon-\bar{X}\|_{\beta,[a,b]}\leq\ &\mathbf{A}_1+K_{\beta,T,|X_0|,f,\sigma}\Psi_1\Lambda_B^{\diamond(\beta)}\delta^\beta+K_{\beta,T}\Psi_1\Phi_\varepsilon+\Psi_2\|\hat{X}^\varepsilon-\bar{X}\|_{\infty,[a,b]}\\&+G^2_{\beta,[a,b]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon)(b-a)^\beta\|\hat{X}^\varepsilon-\bar{X}\|_{\beta,[a,b]}\\&+G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\|\hat{v}^\varepsilon-\bar{v}\|_{2\beta,\Delta_{a,b}},\end{aligned}$$

where we set

$$\begin{aligned}\Psi_1&=1+G^1_{\beta,[0,T]}(\sigma,X^\varepsilon,\hat{X}^\varepsilon,B,v^\varepsilon)+G^2_{\beta,[0,T]}(\sigma,X^\varepsilon,\hat{X}^\varepsilon,B,v^\varepsilon)+G^3_{\beta,[0,T]}(\sigma,\hat{X}^\varepsilon)\|B\|_\beta,\\ \Psi_2&=K(b-a)^{1-\beta}+G^1_{\beta,[0,T]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon)(b-a)^\beta.\end{aligned}$$

Next, by (4.1), we have

$$\begin{aligned}\|\hat{v}^\varepsilon-\bar{v}\|_{2\beta,[a,b]}\leq\ &\sup_{(s,t)\in\Delta}\frac{|\int_s^t\int_s^q(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r\otimes\mathrm{d}B_q|}{(t-s)^{2\beta}}\\&+\sup_{(s,t)\in\Delta_{a,b}}\frac{|\int_s^t\int_s^q(\bar{f}(X^\varepsilon_{r(\delta)})-\bar{f}(X^\varepsilon_r))\mathrm{d}r\otimes\mathrm{d}B_q|}{(t-s)^{2\beta}}\\&+\sup_{(s,t)\in\Delta_{a,b}}\frac{|\int_s^t\int_s^q(\bar{f}(X^\varepsilon_r)-\bar{f}(\hat{X}^\varepsilon_r))\mathrm{d}r\otimes\mathrm{d}B_q|}{(t-s)^{2\beta}}\\&+\sup_{(s,t)\in\Delta_{a,b}}\frac{|\int_s^t\int_s^q(\bar{f}(\hat{X}^\varepsilon_r)-\bar{f}(\bar{X}_r))\mathrm{d}r\otimes\mathrm{d}B_q|}{(t-s)^{2\beta}}\\&+\sup_{(s,t)\in\Delta_{a,b}}\frac{|\int_s^t(\sigma(\hat{X}^\varepsilon_r)-\sigma(\bar{X}_r))\mathrm{d}B^2_{\cdot,t}(r)|}{(t-s)^{2\beta}}\\&+\sup_{(s,t)\in\Delta_{a,b}}\frac{|\int_s^t(\sigma(X^\varepsilon_r)-\sigma(\hat{X}^\varepsilon_r))\mathrm{d}B^2_{\cdot,t}(r)|}{(t-s)^{2\beta}}\\=:&\sum_{j=1}^6\mathbf{B}_j.\end{aligned}$$

Now, we estimate the norm $\|\hat{v}^\varepsilon-\bar{v}\|_{2\beta,[a,b]}$. Let us study $\mathbf{B}_2$, $\mathbf{B}_3$ and $\mathbf{B}_4$. By Fubini's theorem and the argument proposed in [10, p. 2367], it is easy to obtain

$$\sum_{i=2}^4\mathbf{B}_i\leq\|B\|_\beta(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta+K_{\beta,T}\Phi_\varepsilon+K(b-a)^{1-\beta}\|\hat{X}^\varepsilon-\bar{X}\|_{\infty,[a,b]}).$$

Next, thanks to Proposition 2.8 (taking $\lambda=1$), we have

$$\mathbf{B}_5\leq G^4_{\beta,[a,b]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^\beta\|\hat{X}^\varepsilon-\bar{X}\|_{\infty,[a,b]}$$

$$+G^5_{\beta,[a,b]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^\beta\|\hat{X}^\varepsilon-\bar{X}\|_{\beta,[a,b]}$$
$$+G^6_{\beta,[a,b]}(\sigma,\bar{X},B)(b-a)^\beta\|\hat{v}^\varepsilon-\bar{v}\|_{2\beta,\Delta_{a,b}},$$

and by Proposition 2.8 and Lemma 4.2, we have

$$\begin{aligned}\mathbf{B}_6\le{}& G^4_{\beta,[a,b]}(\sigma,X^\varepsilon,\hat{X}^\varepsilon,B,v^\varepsilon,B^2)(b-a)^\beta(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta+\Phi_\varepsilon)\\&+G^5_{\beta,[a,b]}(\sigma,X^\varepsilon,\hat{X}^\varepsilon,B,v^\varepsilon,B^2)(b-a)^\beta(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta+\Phi_\varepsilon)\\&+G^6_{\beta,[a,b]}(\sigma,\hat{X}^\varepsilon,B)(b-a)^\beta\|B\|_\beta(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta+\Phi_\varepsilon).\end{aligned}$$

We take suitable a and b such that

$$G^6_{\beta,[0,T]}(\sigma,\bar{X},B)(b-a)^\beta\le\frac{1}{2}.\tag{4.4}$$

So, we can define Δ^1_β such that $\Delta^1_\beta:=(2G^6_{\beta,[0,T]}(\sigma,\bar{X},B))^{-\frac{1}{\beta}}$. Then, for $(b-a)\le\Delta^1_\beta$ it is easy to obtain

$$\begin{aligned}\|\hat{v}^\varepsilon-\bar{v}\|_{2\beta,\Delta_{a,b}}\le{}&2\mathbf{B}_1+K_{\beta,T,|X_0|,f,\sigma}\Psi_3\Lambda_B^{\diamond(\beta)}\delta^\beta+K_{\beta,T}\Psi_3\Phi_\varepsilon+\Psi_4\|\hat{X}^\varepsilon-\bar{X}\|_{\infty,[a,b]}\\&+2G^5_{\beta,[a,b]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^\beta\|\hat{X}^\varepsilon-\bar{X}\|_{\beta,[a,b]},\end{aligned}$$

where

$$\begin{aligned}\Psi_3={}&2\|B\|_\beta+2G^4_{\beta,[0,T]}(\sigma,X^\varepsilon,\hat{X}^\varepsilon,B,v^\varepsilon,B^2)\\&+2G^5_{\beta,[0,T]}(\sigma,X^\varepsilon,\hat{X}^\varepsilon,B,v^\varepsilon,B^2)+2G^6_{\beta,[0,T]}(\sigma,\hat{X}^\varepsilon,B)\|B\|_\beta,\\\Psi_4={}&2K(b-a)^{1-\beta}\|B\|_\beta+2G^4_{\beta,[0,T]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^\beta.\end{aligned}$$

Next, we have

$$\begin{aligned}\|\hat{X}^\varepsilon-\bar{X}\|_{\beta,[a,b]}\le{}&\mathbf{A}_1+K_{\beta,T,|X_0|,f,\sigma}\Psi_1\Lambda_B^{\diamond(\beta)}\delta^\beta+K_{\beta,T}\Psi_1\Phi_\varepsilon+\Psi_2\|\hat{X}^\varepsilon-\bar{X}\|_{\infty,[a,b]}\\&+G^2_{\beta,[a,b]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^\beta\|\hat{X}^\varepsilon-\bar{X}\|_{\beta,[a,b]}\\&+G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\\&\quad\times[2\mathbf{B}_1+K_{\beta,T,|X_0|,f,\sigma}\Psi_3\Lambda_B^{\diamond(\beta)}\delta^\beta+K_{\beta,T}\Psi_3\Phi_\varepsilon+\Psi_4\|\hat{X}^\varepsilon-\bar{X}\|_{\infty,[a,b]}\\&\quad+2G^5_{\beta,[a,b]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^\beta\|\hat{X}^\varepsilon-\bar{X}\|_{\beta,[a,b]}].\end{aligned}$$

Similar to the definition of Δ^1_β, we take suitable a and b again such that

$$\begin{aligned}&G^2_{\beta,[0,T]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^\beta\\&\quad+2G^5_{\beta,[0,T]}(\sigma,\hat{X}^\varepsilon,\bar{X},B,\hat{v}^\varepsilon,B^2)(b-a)^{2\beta}G^3_{\beta,[0,T]}(\sigma,\bar{X})\le\frac{1}{2}.\end{aligned}\tag{4.5}$$

Then, we have

$$\begin{aligned}
\|\hat{X}^\varepsilon - \bar{X}\|_{\beta,[a,b]} &\le 2\mathbf{A}_1 + 2K_{\beta,T,|X_0|,f,\sigma}\Psi_1\Lambda_B^{\diamond(\beta)}\delta^\beta + 2K_{\beta,T}\Psi_1\Phi_\varepsilon + 2\Psi_2\|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]} \\
&\quad + 2G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta[2\mathbf{B}_1 + K_{\beta,T,|X_0|,f,\sigma}\Psi_3\Lambda_B^{\diamond(\beta)}\delta^\beta + K_{\beta,T}\Psi_3\Phi_\varepsilon \\
&\quad + \Psi_4\|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]}] \\
&= 2\mathbf{A}_1 + 4G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\mathbf{B}_1 \\
&\quad + (2\Psi_1 + 2G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\Psi_3)(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + K_{\beta,T}\Phi_\varepsilon) \\
&\quad + (2\Psi_2 + 2G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\Psi_4)\|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]}.
\end{aligned}$$

Putting

$$\|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]} \le |\hat{X}^\varepsilon_a - \bar{X}_a| + (b-a)^\beta\|\hat{X}^\varepsilon - \bar{X}\|_{\beta,[a,b]},$$

in above equation we have

$$\begin{aligned}
\|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]} &\le |\hat{X}^\varepsilon_a - \bar{X}_a| + (b-a)^\beta\big[2\mathbf{A}_1 + 4G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\mathbf{B}_1 \\
&\quad + (2\Psi_1 + 2G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\Psi_3)(K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}\delta^\beta + K_{\beta,T}\Phi_\varepsilon) \\
&\quad + (2\Psi_2 + 2G^3_{\beta,[a,b]}(\sigma,\bar{X})(b-a)^\beta\Psi_4)\|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]}\big].
\end{aligned}$$

Similar to the definition of Δ^1_β, we take suitable a and b again such that

$$(b-a)^\beta(2\Psi_2 + 2G^3_{\beta,[0,T]}(\sigma,\bar{X})(b-a)^\beta\Psi_4) \le \frac{1}{2}, \tag{4.6}$$

and by Lemma 2.6 and Lemma 4.1, it is easy to know

$$\|\hat{X}^\varepsilon\|_\infty + \|\hat{X}^\varepsilon\|_\beta + \|\hat{v}^\varepsilon\|_{2\beta,\Delta} + \|\bar{X}\|_\infty + \|\bar{X}\|_\beta + \|\bar{v}\|_{2\beta,\Delta} \le K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}$$

holds. Then, we have

$$\|\hat{X}^\varepsilon - \bar{X}\|_{\infty,[a,b]} \le 2|\hat{X}^\varepsilon_a - \bar{X}_a| + K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}(\delta^\beta + \mathbf{A}_1 + \mathbf{B}_1 + \Phi_\varepsilon).$$

Hence,

$$\sup_{0\le t\le b}|\hat{X}^\varepsilon_t - \bar{X}_t| \le 2\sup_{0\le t\le a}|\hat{X}^\varepsilon_t - \bar{X}_t| + K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}(\delta^\beta + \mathbf{A}_1 + \mathbf{B}_1 + \Phi_\varepsilon). \tag{4.7}$$

Now, we can take suitable a and b. There exists Δ^{max}_β such that all a, b with $(b-a) \le \Delta^{max}_\beta$ fulfill (4.4), (4.5) and (4.6), then, it is clear that (4.7) holds for all a and b such that $(b-a) \le \Delta^{max}_\beta$. Then, choose a certain $M = K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}$, we take a partition $0 = t_0 < t_1 < \cdots < t_M = T$ of the interval $[0,T]$ such that $(t_{i+1} - t_i) \le \Delta^{max}_\beta$. Then,

$$\begin{aligned}
\sup_{0\le t\le t_M = T}|\hat{X}^\varepsilon_t - \bar{X}_t| &\le 2\sup_{0\le t\le t_{M-1}}|\hat{X}^\varepsilon_t - \bar{X}_t| \\
&\quad + K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}(\delta^\beta + \mathbf{A}_1 + \mathbf{B}_1 + \Phi_\varepsilon).
\end{aligned} \tag{4.8}$$

Repeating the process M times we obtain

$$\sup_{0\le t\le T}|\hat{X}_t^{\varepsilon}-\bar{X}_t|\le 2^M|\hat{X}_0^{\varepsilon}-\bar{X}_0|+\left(\sum_{k=0}^{M-1}2^k\right)K_{\beta,T,|X_0|,f,\sigma}\Lambda_B^{\diamond(\beta)}(\delta^{\beta}+\mathbf{A}_1+\mathbf{B}_1+\Phi_{\varepsilon})$$
$$\le K_{\beta,T,|X_0|,f,\sigma}2^{\Lambda_B^{\diamond(\beta)}}\Lambda_B^{\diamond(\beta)}(\delta^{\beta}+\mathbf{A}_1+\mathbf{B}_1+\Phi_{\varepsilon}).$$

Finally, we have

$$\|\hat{X}^{\varepsilon}-\bar{X}\|_{\infty}\le K_{\beta,T,|X_0|,f,\sigma}2^{\Lambda_B^{\diamond(\beta)}}\Lambda_B^{\diamond(\beta)}(\delta^{\beta}+\mathbf{A}_1+\mathbf{B}_1+\Phi_{\varepsilon}).$$

This completed the proof of Lemma 4.4. $\square$

By Theorem 3.3, the first level path of the fast component of RDE (1.3) is an Itô SDE (3.3). Thus, it is easy to derive an upper bound of the supremum norm of the solution Y^{ε}. By [27, Lemma 4.3, Lemma 4.4], the following two lemmas are obtained.

Lemma 4.5. *Suppose that* (H1)-(H4) *hold. Then, we have*

$$\sup_{t\in[0,T]}\mathbb{E}[|Y_t^{\varepsilon}|^2]\le K,$$

where $K>0$ is a constant independent of ε.

Lemma 4.6. *Suppose that* (H1)-(H4) *hold. Then, we have*

$$\sup_{t\in[0,T]}\mathbb{E}[|Y_t^{\varepsilon}-\hat{Y}_t^{\varepsilon}|^2]\le K\delta,$$

where $K>0$ is a constant independent of δ and ε.

4.2. Some estimates on the difference between f and $\bar{f}$

Now, let us study $\mathbf{A}_1$. It is easy to see that

$$\begin{aligned}\mathbf{A}_1\le&\sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t(f(X_{r(\delta)}^{\varepsilon},\hat{Y}_r^{\varepsilon})-\bar{f}(X_{r(\delta)}^{\varepsilon}))\mathrm{d}r|}{(t-s)^{\beta}}\mathbf{1}_{\ell}\right\}\\&+\sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t(f(X_{r(\delta)}^{\varepsilon},\hat{Y}_r^{\varepsilon})-\bar{f}(X_{r(\delta)}^{\varepsilon}))\mathrm{d}r|}{(t-s)^{\beta}}\mathbf{1}_{\ell^c}\right\}\\=:&\,\mathbf{A}_{11}+\mathbf{A}_{12},\end{aligned}$$

where $\mathbf{1}_{\cdot}$ is an indicator function, $\ell:=\{t<(\lfloor s\delta^{-1}\rfloor+2)\delta\}$ and $\ell^c:=\{t\ge(\lfloor s\delta^{-1}\rfloor+2)\delta\}$.

On the one hand, by (H3) and the fact that $t-s<\lfloor s\delta^{-1}\rfloor\delta - s + 2\delta \le 2\delta$, we have

$$\begin{aligned}\mathbb{E}[\mathbf{A}_{11}^2] &\le \mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{\frac{|\int_s^t (f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r|^2}{(t-s)^{2\beta}}\mathbf{1}_\ell\Bigg\}\Bigg]\\ &\le K_{\beta,T}\delta.\end{aligned}$$

On the other hand, by (H3) and the fact that $\lfloor\lambda_1\rfloor-\lfloor\lambda_2\rfloor\le\lambda_1-\lambda_2+1$, for $\lambda_1\ge\lambda_2\ge 0$, we have

$$\begin{aligned}\mathbb{E}[\mathbf{A}_{12}^2] &\le K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{\frac{|\int_s^{(\lfloor s\delta^{-1}\rfloor+1)\delta} (f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r|^2}{(t-s)^{2\beta}}\mathbf{1}_{\ell^c}\Bigg\}\Bigg]\\
&\quad+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{\frac{|\int_{\lfloor t\delta^{-1}\rfloor\delta}^{t} (f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r|^2}{(t-s)^{2\beta}}\mathbf{1}_{\ell^c}\Bigg\}\Bigg]\\
&\quad+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{\frac{|\int_{(\lfloor s\delta^{-1}\rfloor+1)\delta}^{\lfloor t\delta^{-1}\rfloor\delta} (f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r|^2}{(t-s)^{2\beta}}\mathbf{1}_{\ell^c}\Bigg\}\Bigg]\\
&\le K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{(t-s)^{1-2\beta}\Bigg|\int_s^{(\lfloor s\delta^{-1}\rfloor+1)\delta}(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r\Bigg|\mathbf{1}_{\ell^c}\Bigg\}\Bigg]\\
&\quad+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{(t-s)^{1-2\beta}\Bigg|\int_{\lfloor t\delta^{-1}\rfloor\delta}^{t}(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r\Bigg|\mathbf{1}_{\ell^c}\Bigg\}\Bigg]\\
&\quad+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{(t-s)^{-2\beta}\Bigg|\sum_{k=\lfloor s\delta^{-1}\rfloor+1}^{\lfloor t\delta^{-1}\rfloor-1}\int_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^2\mathbf{1}_{\ell^c}\Bigg\}\Bigg]\\
&\le K_{\beta,T}\delta+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{(\lfloor t\delta^{-1}\rfloor-\lfloor s\delta^{-1}\rfloor-1)(t-s)^{-2\beta}\\
&\qquad\times\sum_{k=\lfloor s\delta^{-1}\rfloor+1}^{\lfloor t\delta^{-1}\rfloor-1}\Bigg|\int_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^2\mathbf{1}_{\ell^c}\Bigg\}\Bigg]\\
&\le K_{\beta,T}\delta+K_{\beta,T}\delta^{-1}\mathbb{E}\Bigg[\sum_{k=0}^{\lfloor T\delta^{-1}\rfloor-1}\Bigg|\int_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^2\Bigg]\\
&\le K_{\beta,T}\delta+K_{\beta,T}\delta^{-2}\max_{0\le k\le\lfloor T\delta^{-1}\rfloor-1}\mathbb{E}\Bigg[\Bigg|\int_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^2\Bigg].\end{aligned}$$

Now, by the construction of $\hat{Y}^\varepsilon$ and a time shift transformation, for any fixed k and $s\in[0,\delta]$, we have

$$\hat{Y}^{\varepsilon}_{s+k\delta} = \hat{Y}^{\varepsilon}_{k\delta} + \frac{1}{\varepsilon}\int_{k\delta}^{k\delta+s} \tilde{g}(X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_r)\mathrm{d}r + \frac{1}{\sqrt{\varepsilon}}\int_{k\delta}^{k\delta+s} h(X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_r)\mathrm{d}^{\mathrm{I}} W_r$$
$$= \hat{Y}^{\varepsilon}_{k\delta} + \frac{1}{\varepsilon}\int_0^s \tilde{g}\big(X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{r+k\delta}\big)\mathrm{d}r + \frac{1}{\sqrt{\varepsilon}}\int_0^s h\big(X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{r+k\delta}\big)\mathrm{d}^{\mathrm{I}} W^*_r,$$

where $W^*_t = W_{t+k\delta} - W_{k\delta}$ is the shift version of W_t, and hence they have the same distribution.

Let $\bar{W}$ be a Bm and independent of W. Construct a process $Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}$ by means of

$$Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_{s/\varepsilon} = \hat{Y}^{\varepsilon}_{k\delta} + \int_0^{s/\varepsilon} \tilde{g}\big(X^{\varepsilon}_{k\delta}, Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_r\big)\mathrm{d}r + \int_0^{s/\varepsilon} h\big(X^{\varepsilon}_{k\delta}, Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_r\big)\mathrm{d}^{\mathrm{I}}\bar{W}_r$$
$$= \hat{Y}^{\varepsilon}_{k\delta} + \frac{1}{\varepsilon}\int_0^s \tilde{g}\big(X^{\varepsilon}_{k\delta}, Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_{r/\varepsilon}\big)\mathrm{d}r + \frac{1}{\sqrt{\varepsilon}}\int_0^s h\big(X^{\varepsilon}_{k\delta}, Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_{r/\varepsilon}\big)\mathrm{d}^{\mathrm{I}}\bar{\bar{W}}^{\varepsilon}_r, \tag{4.9}$$

where $\bar{\bar{W}}^{\varepsilon}_t = \sqrt{\varepsilon}\bar{W}_{t/\varepsilon}$ is the scaled version of $\bar{W}_t$. Because both W^* and $\bar{\bar{W}}$ are independent of $(X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta})$, by comparison, yields

$$\big(X^{\varepsilon}_{k\delta}, \{\hat{Y}^{\varepsilon}_{s+k\delta}\}_{s\in[0,\delta)}\big) \sim \big(X^{\varepsilon}_{k\delta}, \{Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_{s/\varepsilon}\}_{s\in[0,\delta)}\big), \tag{4.10}$$

where $\sim$ denotes coincidence in distribution sense. Thus, we have

$$\mathbb{E}[\mathbf{A}_1^2] \le K_{\beta,T}\delta + K_{\beta,T}\delta^{-2} \max_{0\le k\le \lfloor T\delta^{-1}\rfloor -1} \mathbb{E}\Bigg[\bigg|\int_{k\delta}^{(k+1)\delta} (f(X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_r) - \bar{f}(X^{\varepsilon}_{k\delta}))\mathrm{d}r\bigg|^2\Bigg]$$
$$\le K_{\beta,T}\delta + K_{\beta,T}\varepsilon^2\delta^{-2} \max_{0\le k\le \lfloor T\delta^{-1}\rfloor -1} \int_0^{\frac{\delta}{\varepsilon}}\int_{\theta}^{\frac{\delta}{\varepsilon}} \mathcal{J}_k(s,\theta)\mathrm{d}s\mathrm{d}\theta,$$

where

$$\mathcal{J}_k(s,\theta) = \mathbb{E}[\langle f(X^{\varepsilon}_{k\delta}, Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_s) - \bar{f}(X^{\varepsilon}_{k\delta}), f(X^{\varepsilon}_{k\delta}, Y^{X^{\varepsilon}_{k\delta}, \hat{Y}^{\varepsilon}_{k\delta}}_{\theta}) - \bar{f}(X^{\varepsilon}_{k\delta})\rangle].$$

Through the argument as in Appendix A (the similar argument could also be found in [27, Appendix B]), i.e., for any $0 \le \theta \le s \le \frac{\delta}{\varepsilon}$ and $k = 0, 1, \dots, [T/\delta] - 1$, we have

$$\mathcal{J}_k(s,\theta) \le K e^{-\frac{\beta_1}{2}(s-\theta)}\mathbb{E}[(1 + |X^{\varepsilon}_{k\delta}|^2 + |\hat{Y}^{\varepsilon}_{k\delta}|^2)] \le K_{T,|X_0|,|Y_0|} e^{-\frac{\beta_1}{2}(s-\theta)},$$

where β_1 is defined in (H4). Here, Lemmas 4.1, 4.5 and 4.6 were used for the last inequality.

Thus, we have

$$\mathbb{E}[\mathbf{A}_1^2] \leq K_{\beta,T}(\delta + \varepsilon\delta^{-1}). \tag{4.11}$$

For the term $\mathbf{B}_1$, by Fubini's theorem and the argument proposed in [10, p. 2367], we have

$$\begin{aligned}
\mathbf{B}_1 &= \sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t\int_s^q(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\mathrm{d}r\otimes\mathrm{d}B_q|}{(t-s)^{2\beta}}\right\}\\
&\leq \sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_r^t\mathrm{d}B_q\mathrm{d}r|}{(t-s)^{2\beta}}\right\}\\
&\leq \sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_{r(\delta)}^t\mathrm{d}B_q\mathrm{d}r|}{(t-s)^{2\beta}}\right\}\\
&\quad+\sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_r^{r(\delta)}\mathrm{d}B_q\mathrm{d}r|}{(t-s)^{2\beta}}\right\}\\
&=: \mathbf{B}_{11}+\mathbf{B}_{12}.
\end{aligned}$$

Let us study $\mathbf{B}_{11}$. Similarly, we have

$$\begin{aligned}
\mathbb{E}[\mathbf{B}_{11}^2] &\leq \mathbb{E}\left[\sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_{r(\delta)}^t\mathrm{d}B_q\mathrm{d}r|^2}{(t-s)^{4\beta}}\mathbf{1}_\ell\right\}\right]\\
&\quad+\mathbb{E}\left[\sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^t(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_{r(\delta)}^t\mathrm{d}B_q\mathrm{d}r|^2}{(t-s)^{4\beta}}\mathbf{1}_{\ell^c}\right\}\right]\\
&\leq K_{\beta,T}\mathbb{E}[\|B\|_\beta^2]\delta\\
&\quad+K\mathbb{E}\left[\sup_{(s,t)\in\Delta}\left\{\frac{|\int_s^{(\lfloor s\delta^{-1}\rfloor+1)\delta}(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_{r(\delta)}^t\mathrm{d}B_q\mathrm{d}r|^2}{(t-s)^{4\beta}}\mathbf{1}_{\ell^c}\right\}\right]\\
&\quad+K\mathbb{E}\left[\sup_{(s,t)\in\Delta}\left\{\frac{|\int_{\lfloor t\delta^{-1}\rfloor\delta}^t(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_{r(\delta)}^t\mathrm{d}B_q\mathrm{d}r|^2}{(t-s)^{4\beta}}\mathbf{1}_{\ell^c}\right\}\right]\\
&\quad+K\mathbb{E}\left[\sup_{(s,t)\in\Delta}\left\{\frac{|\int_{(\lfloor s\delta^{-1}\rfloor+1)\delta}^{\lfloor t\delta^{-1}\rfloor\delta}(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_{r(\delta)}^t\mathrm{d}B_q\mathrm{d}r|^2}{(t-s)^{4\beta}}\mathbf{1}_{\ell^c}\right\}\right]\\
&\leq K_{\beta,T}\mathbb{E}[\|B\|_\beta^2]\delta+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{(t-s)^{1-3\beta}\|B\|_\beta\\
&\quad\times\left|\int_s^{(\lfloor s\delta^{-1}\rfloor+2)\delta}(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int_{r(\delta)}^t\mathrm{d}B_q\mathrm{d}r\right|\mathbf{1}_{\ell^c}\Bigg\}\Bigg]
\end{aligned}$$

$$
\begin{aligned}
&+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{(t-s)^{1-3\beta}\|B\|_\beta \\
&\quad\times\Bigg|\int\limits_{\lfloor t\delta^{-1}\rfloor\delta}^{t}(f(X^\varepsilon_{r(\delta)},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{r(\delta)}))\otimes\int\limits_{r(\delta)}^{t}\mathrm{d}B_q\mathrm{d}r\Bigg|\mathbf{1}_{\ell^c}\Bigg\}\Bigg] \\
&+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{\frac{\big|\sum_{\lfloor s\delta^{-1}\rfloor+1}^{\lfloor t\delta^{-1}\rfloor-1}\int_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\otimes\int_{k\delta}^{t}\mathrm{d}B_q\mathrm{d}r\big|^2}{(t-s)^{4\beta}}\mathbf{1}_{\ell^c}\Bigg\}\Bigg] \\
&\le K_{\beta,T}\mathbb{E}[\|B\|_\beta^2]\delta+K\mathbb{E}\Bigg[\sup_{(s,t)\in\Delta}\Bigg\{(\lfloor t\delta^{-1}\rfloor-\lfloor s\delta^{-1}\rfloor-1)(t-s)^{-2\beta}\|B\|_\beta^2 \\
&\quad\times\sum_{k=\lfloor s\delta^{-1}\rfloor+1}^{\lfloor t\delta^{-1}\rfloor-1}\Bigg|\int\limits_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^2\mathbf{1}_{\ell^c}\Bigg\}\Bigg] \\
&\le K_{\beta,T}\mathbb{E}[\|B\|_\beta^2]\delta \\
&\quad+K_{\beta,T}\delta^{-1}\sum_{k=0}^{\lfloor T\delta^{-1}\rfloor-1}\mathbb{E}\Bigg[\|B\|_\beta^2\Bigg|\int\limits_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^2\Bigg] \\
&\le K_{\beta,T}\mathbb{E}[\|B\|_\beta^2]\delta+K_{\beta,T}\delta^{-1}\big(\mathbb{E}\big[\|B\|_\beta^4\big]\big)^{\frac{1}{2}} \\
&\quad\times\sum_{k=0}^{\lfloor T\delta^{-1}\rfloor-1}\Bigg(\mathbb{E}\Bigg[\Bigg|\int\limits_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^4\Bigg]\Bigg)^{\frac{1}{2}} \\
&\le K_{\beta,T}\delta+K_{\beta,T}\delta^{-1}\max_{0\le k\le\lfloor T\delta^{-1}\rfloor-1}\Bigg(\mathbb{E}\Bigg[\Bigg|\int\limits_{k\delta}^{(k+1)\delta}(f(X^\varepsilon_{k\delta},\hat{Y}^\varepsilon_r)-\bar{f}(X^\varepsilon_{k\delta}))\mathrm{d}r\Bigg|^2\Bigg]\Bigg)^{\frac{1}{2}} \\
&\le K_{\beta,T}(\delta+\varepsilon^{\frac{1}{2}}\delta^{-\frac{1}{2}}).
\end{aligned}
$$

To proceed, it is easy to obtain

$$\mathbb{E}[\mathbf{B}_{12}^2]\le K_{\beta,T}\delta^{2\beta}.$$

Thus, we have

$$\mathbb{E}[\mathbf{B}_1^2]\le K_{\beta,T}(\delta^{2\beta}+\varepsilon^{\frac{1}{2}}\delta^{-\frac{1}{2}}). \tag{4.12}$$

Now, let us study $\mathbb{E}[\Phi_\varepsilon^2]$.

$$\mathbb{E}[\Phi_\varepsilon^2]\le 2\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Bigg|\int\limits_0^t(f(X^\varepsilon_{s(\delta)},\hat{Y}^\varepsilon_s)-f(X^\varepsilon_{s(\delta)},Y^\varepsilon_s))\mathrm{d}s\Bigg|^2\Bigg]$$

$$+2\mathbb{E}\Big[\sup_{(s,t)\in\Delta}\frac{\big|\int_s^t(f(X^{\varepsilon}_{r(\delta)},\hat{Y}^{\varepsilon}_r)-f(X^{\varepsilon}_{r(\delta)},Y^{\varepsilon}_r))\mathrm{d}r\big|^2}{(t-s)^{2\beta}}\Big]$$
$$=: 2\mathbf{C}_1+2\mathbf{C}_2.$$

Through a similar argument as in the estimates of $\mathbb{E}[\mathbf{A}_{11}^2]$ and $\mathbb{E}[\mathbf{A}_{12}^2]$ and by Lemma 4.6, we have

$$\begin{aligned}
\mathbf{C}_1 &\le 2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_{\lfloor t\delta^{-1}\rfloor}^{t}(f(X^{\varepsilon}_{s(\delta)},\hat{Y}^{\varepsilon}_s)-f(X^{\varepsilon}_{s(\delta)},Y^{\varepsilon}_s))\mathrm{d}s\Big|^2\Big]\\
&\quad+2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\sum_{k=0}^{\lfloor t\delta^{-1}\rfloor-1}\int_{k\delta}^{(k+1)\delta}(f(X^{\varepsilon}_{k\delta},\hat{Y}^{\varepsilon}_s)-f(X^{\varepsilon}_{k\delta},Y^{\varepsilon}_s))\mathrm{d}s\Big|^2\Big]\\
&\le K\delta^2+K_{\beta,T}\delta^{-2}\max_{0\le k\le\lfloor T\delta^{-1}\rfloor-1}\mathbb{E}\Big[\Big|\int_{k\delta}^{(k+1)\delta}(f(X^{\varepsilon}_{k\delta},\hat{Y}^{\varepsilon}_s)-f(X^{\varepsilon}_{k\delta},Y^{\varepsilon}_s))\mathrm{d}s\Big|^2\Big]\\
&\le K\delta^2+K_{\beta,T}\delta^{-1}\max_{0\le k\le\lfloor T\delta^{-1}\rfloor-1}\int_{k\delta}^{(k+1)\delta}\mathbb{E}[|f(X^{\varepsilon}_{k\delta},\hat{Y}^{\varepsilon}_s)-f(X^{\varepsilon}_{k\delta},Y^{\varepsilon}_s)|^2]\mathrm{d}s\\
&\le K_{\beta,T}\delta,
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{C}_2 &\le K_{\beta,T}\delta+K_{\beta,T}\delta^{-2}\max_{0\le k\le\lfloor T\delta^{-1}\rfloor-1}\mathbb{E}\Big[\Big|\int_{k\delta}^{(k+1)\delta}(f(X^{\varepsilon}_{k\delta},\hat{Y}^{\varepsilon}_s)-f(X^{\varepsilon}_{k\delta},Y^{\varepsilon}_s))\mathrm{d}s\Big|^2\Big]\\
&\le K_{\beta,T}\delta+K_{\beta,T}\delta^{-1}\max_{0\le k\le\lfloor T\delta^{-1}\rfloor-1}\int_{k\delta}^{(k+1)\delta}\mathbb{E}[|f(X^{\varepsilon}_{k\delta},\hat{Y}^{\varepsilon}_s))-f(X^{\varepsilon}_{k\delta},Y^{\varepsilon}_s))|^2]\mathrm{d}s\\
&\le K_{\beta,T}\delta.
\end{aligned}$$

Thus, we have

$$\mathbb{E}[\Phi_{\varepsilon}^2]\le K_{\beta,T}\delta. \tag{4.13}$$

4.3. Proof of Theorem 1.2

By Lemma 4.3 and Lemma 4.4, it is easy to have

$$\begin{aligned}
\|X^{\varepsilon}-\bar{X}\|_{\infty}&\le\|X^{\varepsilon}-\hat{X}^{\varepsilon}\|_{\infty}+\|\hat{X}^{\varepsilon}-\bar{X}\|_{\infty}\\
&\le K_{\beta,T,|X_0|,f,\sigma}2^{\Lambda_B^{\diamond(\beta)}}\Lambda_B^{\diamond(\beta)}(\delta^{\beta}+\mathbf{A}_1+\mathbf{B}_1+\Phi_{\varepsilon}),\quad \text{a.s.}
\end{aligned}$$

Next, for each $R > 1$, set $D := \{\Lambda_B \leq R\}$ and $D^c := \{\Lambda_B > R\}$, then, by (4.11), (4.12) and (4.13), we have

$$\begin{aligned}\mathbb{E}[\|X^\varepsilon - \bar{X}\|_\infty \mathbf{1}_D] &\leq \mathbb{E}[K_{\beta,T,|X_0|,f,\sigma} 2^{\Lambda_B^{\diamond(\beta)}} \Lambda_B^{\diamond(\beta)} (\delta^\beta + \mathbf{A}_1 + \mathbf{B}_1 + \Phi_\varepsilon)] \\ &\leq K_R K_{\beta,T,|X_0|,f,\sigma} (\delta^\beta + \varepsilon^{\frac{1}{4}} \delta^{-\frac{1}{4}}),\end{aligned} \tag{4.14}$$

where $K_R > 0$ is a constant and

$$\begin{aligned}\mathbb{E}[\|X^\varepsilon - \bar{X}\|_\infty \mathbf{1}_{D^c}] &\leq (\mathbb{E}[\|X^\varepsilon - \bar{X}\|_\infty^2])^{\frac{1}{2}} P(\Lambda_B > R)^{\frac{1}{2}} \\ &\leq (\mathbb{E}[\|X^\varepsilon\|_\infty^2 + \|\bar{X}\|_\infty^2])^{\frac{1}{2}} P(\Lambda_B > R)^{\frac{1}{2}} \\ &\leq K_{\beta,T,|X_0|,f,\sigma} (\mathbb{E}[\Lambda_B^{\diamond(\beta)}])^{\frac{1}{2}} P(\Lambda_B > R)^{\frac{1}{2}},\end{aligned} \tag{4.15}$$

where $K_{\beta,T,|X_0|,f,\sigma}$, $\mathbb{E}[\Lambda_B^{\diamond(\beta)}]$ and $P(\Lambda_B > R)$ are all independent of ε.

Putting (4.14) and (4.15) together and choosing $\delta := \varepsilon\sqrt{-\ln \varepsilon}$, we have

$$\limsup_{\varepsilon \to 0} \mathbb{E}[\|X^\varepsilon - \bar{X}\|_\infty] \leq K P(\Lambda_B > R)^{\frac{1}{2}},$$

where $K > 0$ is a constant which is independent of ε and R. Then, let $R \to \infty$, we have

$$\lim_{\varepsilon \to 0} \mathbb{E}[\|X^\varepsilon - \bar{X}\|_\infty] = 0.$$

Thus, the statement of Theorem 1.2 is obtained.

Acknowledgment

The authors are grateful to Professor Yu Ito for stimulating discussions during the preparation of this work. They also thank Professors David Nualart and Yaozhong Hu for helpful comments. B. Pei would like to thank JSPS for Postdoctoral Fellowships for Research in Japan (Standard).

B. Pei was partially supported by the National Natural Science Foundation of China under Grant No. 11802216, the Fundamental Research Funds for the Central Universities, the Young Talent fund of University Association for Science and Technology in Shaanxi, China, and JSPS Grant-in-Aid for JSPS Fellows under Grant No. JP18F18314. Y. Inahama was partially supported by JSPS KAKENHI under Grant No. JP20H01807 and Grant-in-Aid for JSPS Fellows under Grant No. JP18F18314. Y. Xu was partially supported by the National Natural Science Foundation of China under Grant No. 12072264, the Research Funds for Interdisciplinary Subject of Northwestern Polytechnical University, the Shaanxi Provincial Key R&D Program under Grant No. 2019TD-010 and No. 2020KW-013.

Appendix A

Let $\bar{W}$ be as in (4.9) and $\mathbb{Q}^\phi$ denote the probability law of the diffusion process $\{Y_t^\xi\}_{t\geq 0}$ which is governed by following equation

$$\mathrm{d}Y_t^\xi = \tilde{g}(\xi, Y_t^\xi)\mathrm{d}t + h(\xi, Y_t^\xi)\mathrm{d}^{\mathrm{I}}\bar{W}_t,$$

with initial value $Y_0^\xi = \phi$ and we denote the solution by $\{Y_t^{\xi,\phi}\}_{t\geq 0}$. The expectation with respect to $\mathbb{Q}^\phi$ is denoted by $\mathbb{E}^\phi$. Hence, we have $\mathbb{E}^\phi[\Psi(Y_t^\xi)] = \mathbb{E}[\Psi(Y_t^{\xi,\phi})]$, for all bounded function Ψ. For more details on $\mathbb{Q}^\phi$, the readers are referred to [25, p. 110]. Let $\mathcal{F}_t^\xi$ be the σ-field generated by $\{Y_r^{\xi,\phi}, r \leq t\}$ and set

$$\mathcal{J}_k(s,\theta,\xi,\phi) := \mathbb{E}[\langle f(\xi, Y_s^{\xi,\phi}) - \bar{f}(\xi), f(\xi, Y_\theta^{\xi,\phi}) - \bar{f}(\xi)\rangle].$$

Then, we have

$$\begin{aligned}\mathcal{J}_k(s,\theta,\xi,\phi) &= \mathbb{E}^\phi[\langle f(\xi, Y_s^\xi) - \bar{f}(\xi), f(\xi, Y_\theta^\xi) - \bar{f}(\xi)\rangle] \\ &= \mathbb{E}^\phi[\mathbb{E}^\phi[\langle f(\xi, Y_s^\xi) - \bar{f}(\xi), f(\xi, Y_\theta^\xi) - \bar{f}(\xi)\rangle|\mathcal{F}_\theta^\xi]] \\ &= \mathbb{E}^\phi[\langle f(\xi, Y_\theta^\xi) - \bar{f}(\xi), \mathbb{E}^\phi[(f(\xi, Y_s^\xi) - \bar{f}(\xi))|\mathcal{F}_\theta^\xi]\rangle].\end{aligned}$$

To proceed, by invoking the Markov property of $\{Y_t^{\xi,\phi}\}_{t\geq 0}$, we have

$$\mathcal{J}_k(s,\theta,\xi,\phi) = \mathbb{E}^\phi[\langle f(\xi, Y_\theta^\xi) - \bar{f}(\xi), \mathbb{E}^{Y_\theta^{\xi,\phi}}[f(\xi, Y_{s-\theta}^\xi) - \bar{f}(\xi)]\rangle],$$

where $\mathbb{E}^{Y_\theta^{\xi,\phi}}[f(\xi, Y_{s-\theta}^\xi) - \bar{f}(\xi)]$ means the function $\mathbb{E}^\phi[f(\xi, Y_{s-\varsigma}^\xi) - \bar{f}(\xi)]$ evaluated at $\phi = Y_\theta^{\xi,\phi}$.

Using Hölder's inequality and the boundedness of the function f, we obtain

$$\mathcal{J}_k(s,\theta,\xi,\phi) \leq K(\mathbb{E}^\phi[|f(\xi, Y_\theta^\xi) - \bar{f}(\xi)|^2])^{\frac{1}{2}}(\mathbb{E}^\phi[|\mathbb{E}^{Y_\theta^{\xi,\phi}}[f(\xi, Y_{s-\theta}^\xi) - \bar{f}(\xi)]|^2])^{\frac{1}{2}}.$$

In view of [27, Lemma 0.10], we have

$$\mathcal{J}_k(s,\theta,\xi,\phi) \leq K(1 + |\xi|^2 + |\phi|^2)e^{-\frac{\beta_1}{2}(s-\theta)}. \tag{A.1}$$

Let $\mathcal{M}_{k\delta}^\varepsilon$ be the σ-field generated by $X_{k\delta}^\varepsilon$ and $\hat{Y}_{k\delta}^\varepsilon$ that is independent of $\{Y_t^{\xi,\phi}\}_{t\geq 0}$. By adopting the approach in [25, Theorem 7.1.2]. We can show

$$\begin{aligned}\mathcal{J}_k(s,\theta) &= \mathbb{E}[\mathbb{E}[\langle f(X_{k\delta}^\varepsilon, Y_s^{X_{k\delta}^\varepsilon, \hat{Y}_{k\delta}^\varepsilon}) - \bar{f}(X_{k\delta}^\varepsilon), f(X_{k\delta}^\varepsilon, Y_\theta^{X_{k\delta}^\varepsilon, \hat{Y}_{k\delta}^\varepsilon}) - \bar{f}(X_{k\delta}^\varepsilon)\rangle|\mathcal{M}_{k\delta}^\varepsilon]] \\ &= \mathbb{E}[\mathcal{J}_k(s,\theta,\xi,\phi)|_{(\xi,\phi)=(X_{k\delta}^\varepsilon, \hat{Y}_{k\delta}^\varepsilon)}],\end{aligned}$$

which, with the aid of (A.1), yields

$$\mathcal{J}_k(s,\theta) \leq K\mathbb{E}[(1 + |X_{k\delta}^\varepsilon|^2 + |\hat{Y}_{k\delta}^\varepsilon|^2)]e^{-\frac{\beta_1}{2}(s-\theta)}.$$

This completes the proof of the claim. □

References

[1] M. Besalú, G. Binotto, C. Rovira, Convergence of delay equations driven by a Hölder continuous function of order $\beta \in (\frac{1}{3}, \frac{1}{2})$, Electron. J. Differ. Equ. 65 (2020) 1–27.

[2] M. Besalú, D. Márquez-Carreras, C. Rovira, Delay equations with non-negativity constraints driven by a Hölder continuous function of order $\beta \in (\frac{1}{3}, \frac{1}{2})$, Potential Anal. 41 (2014) 117–141.

[3] M. Besalú, D. Nualart, Estimates for the solution to stochastic differential equations driven by a fractional Brownian motion with Hurst parameter $H \in (\frac{1}{3}, \frac{1}{2})$, Stoch. Dyn. 11 (2011) 243–263.

[4] L. Coutin, Z. Qian, Stochastic analysis, rough path analysis and fractional Brownian motions, Probab. Theory Relat. Fields 122 (2002) 108–140.

[5] J. Diehl, H. Oberhauser, S. Riedel, A Lévy area between Brownian motion and rough paths with applications to robust nonlinear filtering and rough partial differential equations, Stoch. Process. Appl. 125 (2015) 161–181.

[6] J. Duan, W. Wang, Effective Dynamics of Stochastic Partial Differential Equations, Elsevier, 2014.

[7] M. Freidlin, A. Wentzell, Random Perturbations of Dynamical Systems, Springer, 2012.

[8] P.K. Friz, N.B. Victoir, Multidimensional Stochastic Processes as Rough Paths: Theory and Applications, Cambridge University Press, 2010.

[9] P.K. Friz, M. Hairer, A Course on Rough Paths. With an Introduction to Regularity Structures, Springer, 2014.

[10] M. Garrido-Atienza, B. Schmalfuss, Local stability of differential equations driven by Hölder-continuous paths with Hölder index in $(\frac{1}{3}, \frac{1}{2})$, SIAM J. Appl. Dyn. Syst. 17 (2018) 2352–2380.

[11] D. Givon, Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems, Multiscale Model. Simul. 6 (2007) 577–594.

[12] M. Gubinelli, Controlling rough paths, J. Funct. Anal. 216 (2004) 86–140.

[13] J. Guerra, D. Nualart, Stochastic differential equations driven by fractional Brownian motion and standard Brownian motion, Stoch. Anal. Appl. 26 (2008) 1053–1075.

[14] M. Hairer, X.-M. Li, Averaging dynamics driven by fractional Brownian motion, Ann. Probab. 48 (2020) 1826–1860.

[15] Y. Hu, D. Nualart, Rough path analysis via fractional calculus, Trans. Am. Math. Soc. 361 (2009) 2689–2718.

[16] Y. Ito, A fractional calculus approach to rough integration, Kyoto J. Math. 59 (2019) 553–573.

[17] R. Khasminskii, On an averaging principle for Itô stochastic differential equations, Kibernetika 4 (1968) 260–279.

[18] K. Kubilius, The existence and uniqueness of the solution of an integral equation driven by a p-semimartingale of special type, Stoch. Process. Appl. 98 (2002) 289–315.

[19] W. Liu, M. Röckner, X. Sun, Y. Xie, Averaging principle for slow-fast stochastic differential equations with time dependent locally Lipschitz coefficients, J. Differ. Equ. 268 (2020) 2910–2948.

[20] T. Lyons, Differential equations driven by rough signals (I): an extension of an inequality of L. C. Young, Math. Res. Lett. 1 (4) (1994) 451–464.

[21] T. Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoam. 14 (2) (1998) 215–310.

[22] T. Lyons, Z. Qian, System Control and Rough Paths, Oxford University Press, 2002.

[23] T. Lyons, M. Caruana, T. Lévy, Differential Equations Driven by Rough Paths, Springer Berlin Heidelberg, 2007.

[24] A. Neuenkirch, T. Shalaiko, The relation between mixed and rough SDEs and its application to numerical methods, Stoch. Anal. Appl. 33 (2015) 927–942.

[25] B. Øksendal, Stochastic Differential Equations, Springer, Heidelberg, 2003.

[26] D. Nualart, A. Răşcanu, Differential equations driven by fractional Brownian motion, Collect. Math. 53 (2002) 55–81.

[27] B. Pei, Y. Inahama, Y. Xu, Averaging principles for mixed fast-slow systems driven by fractional Brownian motion, arXiv:2001.06945, 2020.

[28] B. Pei, Y. Inahama, Y. Xu, Pathwise unique global solutions and stochastic averaging for mixed stochastic partial differential equations driven by fractional Brownian motion and Brownian motion, arXiv:2004.05305, 2020.

[29] B. Pei, Y. Xu, J.-L. Wu, Stochastic averaging for stochastic differential equations driven by fractional Brownian motion and standard Brownian motion, Appl. Math. Lett. 100 (2020) 106006.

[30] B. Pei, Y. Xu, J.-L. Wu, Two-time-scales hyperbolic-parabolic equations driven by Poisson random measures: Existence, uniqueness and averaging principles, J. Math. Anal. Appl. 447 (1) (2017) 243–268.

[31] S. Riedel, M. Scheutzow, Rough differential equations with unbounded drift term, J. Differ. Equ. 262 (2017) 283–312.

[32] R. Wang, Y. Xu, Stochastic averaging for the non-autonomous mixed stochastic differential equations with locally Lipschitz coefficients, arXiv:2009.05421, 2020.

[33] Y. Xu, J. Duan, W. Xu, An averaging principle for stochastic dynamical systems with Lévy noise, Phys. D: Nonlinear Phenom. 240 (2011) 1395–1401.
[34] Y. Xu, R. Guo, D. Liu, H. Zhang, J. Duan, Stochastic averaging principle for dynamical systems with fractional Brownian motion, Discrete Contin. Dyn. Syst., Ser. B 19 (2014) 1197–1212.
[35] Y. Xu, B. Pei, R. Guo, Stochastic averaging for slow-fast dynamical systems with fractional Brownian motion, Discrete Contin. Dyn. Syst., Ser. B 20 (2015) 2257–2267.
[36] Y. Xu, B. Pei, J.-L. Wu, Stochastic averaging principle for differential equations with non-Lipschitz coefficients driven by fractional Brownian motion, Stoch. Dyn. 17 (2017) 1750013.
[37] M. Zähle, Integration with respect to fractal functions and stochastic calculus I, Probab. Theory Relat. Fields 111 (1998) 333–374.